

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 7, 1999, 21:20:10 ; Search time 465.18 Seconds

(without alignments)
11875.397 Million cell updates/sec

Title: US-09-104-063-3

Perfect score: 1737

Sequence: 1 GAATTCAGACTGCTGCTGGCGG.....CCGCCACAGACTGGAATTC 1737

Scoring table: OLIGO_NUC

Searched: 579419 seqs, 1590154680 residues

Database : GenEmbl :
1: gb_ba1 :
2: gb_ba2 :
3: gb_om :
4: gb_ov :
5: gb_pat :
6: gb_ph :
7: gb_pl1 :
8: gb_pl2 :
9: gb_pl3 :
10: gb_pl4 :
11: gb_pl5 :
12: gb_pl6 :
13: gb_pl7 :
14: gb_pl8 :
15: gb_pl9 :
16: gb_pl10 :
17: gb_pl11 :
18: gb_pl12 :
19: gb_pl13 :
20: gb_pl14 :
21: gb_pl15 :
22: gb_pl16 :
23: gb_pl17 :
24: gb_pl18 :
25: gb_pl19 :
26: gb_pl20 :
27: gb_pl21 :
28: gb_pl22 :
29: gb_pl23 :
30: gb_pl24 :
31: gb_pl25 :
32: gb_pl26 :
33: gb_pl27 :
34: gb_pl28 :
35: gb_pl29 :
36: gb_pl30 :
37: gb_pl31 :
38: gb_pl32 :
39: gb_pl33 :
40: gb_pl34 :
41: gb_pl35 :
42: gb_pl36 :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1737	100.0	1737	5	AR015970	AR015970 Sequence

2	1737	100.0	1737	5	I13753	I13753 Sequence 4
3	1737	100.0	1737	5	I24455	I24455 Sequence 3
4	1642	94.5	1645	10	HSNPRFLA	X71635 H. sapiens m
5	1601	92.2	1664	9	HUMHMR89	D10924 Human mRNA
6	1601	92.2	1664	14	G28514	G28514 human STS S
7	1564	90.0	5161	11	AF005058	AF005058 Homo sapi
8	1536	88.4	3733	10	HSXCR4	Y14739 Homo sapien
9	1513	87.1	8747	9	HSR24859	AJ224860 Homo sapi
10	1513	87.1	8747	11	AF032572	AF032572 Homo sapi
11	1278	73.6	1637	9	HUMSTSR	M99293 Homo sapien
12	1245	71.7	1670	9	HUMGPCR	L06797 Human (clon
13	1108	63.8	1225	9	HUMNYRCA	L01639 Human (clon
14	1059	61.0	1059	11	AF025375	AF025375 Homo sapi
15	985	56.7	1087	11	PTU89798	AF051906 Cercopit
16	341	19.6	1059	10	MM093311	U93311 Macaca mula
17	341	19.6	1087	11	AF001928	AF001928 Macaca mu
18	336	19.3	1078	9	D86579	D86579 Macaca fasc
19	313	18.0	1126	11	AF019378	AF019378 Cercopit
20	313	18.0	1059	11	AF031089	AF031089 Papio ham
21	239	13.8	1059	11	AF051906	AF051906 Cercopit
22	239	13.8	1059	11	MM073740	U73740 Macaca mula
23	131	7.5	1088	9	AB015943	AB015943 Chirocep
24	77	4.4	1689	3	FCA9816	AJ009816 Felis cat
25	77	4.4	1170	3	FCU63558	U63558 Felis catus
26	77	4.4	1140	3	FCU92795	U92795 Felis catus
27	70	4.0	95	9	HUMNYR82	D28433 Human mRNA
28	62	3.6	1280	3	BOVNYR	M86739 Cow neurope
29	56	3.2	1223	12	MM059760	U59760 Mus musculu
30	56	3.2	3366	12	MM065580	U65580 Mus musculu
31	56	3.2	1362	12	RNU54791	U54791 Rattus norv
32	56	3.2	1050	12	RNU90610	U90610 Rattus norv
33	53	3.1	3581	5	E07381	E07381 CDNA encodi
34	53	3.1	995	42	AF002985	AF002985 Homo sapi
35	50	2.6	578	3	OMI38942	U38942 Ovis aries
36	50	2.9	60	5	I88893	I88893 Sequence 11
37	50	2.9	1877	12	D87747	D87747 Mus musculu
38	50	2.9	495	36	S81785	S81785 Myr p II-ma
39	49	2.8	3362	4	AF139986	AF139986 Xenopus l
40	49	2.8	2035	4	AF153205	AF153205 Gallus ga
41	49	2.8	1772	7	SACATATAS	Z96633 Solanella
42	49	2.8	711	8	AF069331	AF069331 Hordeum v
43	49	2.8	1207	8	AF094831	AF094831 Zantedesc
44	49	2.8	1845	10	HSM800467	AL050172 Homo sapi
45	49	2.8	1403	12	MMU132433	AJ132433 Mus muscu
46	49	2.8	754	12	MUSCALTRA	D16301 Mouse mRNA
47	49	2.8	987	36	AFCAATP	X76691 A. francisca
48	49	2.8	982	36	BMOTELRPT	D13554 Bombyx mori
49	48	2.8	1032	3	AF013214	AF013214 Bos tauru
50	48	2.8	1338	4	ECCD44	X66862 E. caballus
51	48	2.8	2374	4	XLA5870	AJ005870 Xenopus l
52	48	2.8	1490	5	A41579	A41579 Sequence 5
53	48	2.8	1474	5	A65341	A65341 Sequence 64
54	48	2.8	1474	5	E06743	E06743 CDNA encodi
55	48	2.8	2638	7	ATRAU0058	AJ000058 Arabidops
56	48	2.8	1425	7	ATGOX2353	X83380 A. thaliana
57	48	2.8	1083	7	ATH10467	AJ010467 Arabidops
58	48	2.8	1586	7	BNPML3	X73850 B. napus (pN
59	48	2.8	1493	7	HVCHS2	Y09233 H. vulgare m
60	48	2.8	1572	8	AF014053	AF014053 Nicotiana
61	48	2.8	2047	8	AF035815	AF035815 Zea mays
62	48	2.8	1960	8	AF150881	AF150881 Lycopersi
63	48	2.8	893	8	ATU75199	U75199 Arabidopsi
64	48	2.8	1589	8	PTX13769	Y13769 Populus tri
65	48	2.8	5503	10	HSCH16FAA	X99226 H. sapiens m
66	48	2.8	1887	10	HSGLCNACT	Z82022 H. sapiens m
67	48	2.8	3319	10	HSNRRP1	X66975 H. sapiens m
68	48	2.8	1746	10	HSR800208	AL049423 Homo sapi
69	48	2.8	3462	10	HSR800550	AL050553 Homo sapi
70	48	2.8	1373	10	HSY14040	Y14040 Homo sapien
71	48	2.8	185741	11	AC006203	AC006203 Homo sapi
72	48	2.8	1448	12	AF067728	AF067728 Rattus no
73	48	2.8	1095	12	CATRANS	X98066 C. aureus mR
74	48	2.8	159319	34	HS40E16	AL031963 Homo sapi

75	48	2.8	944	36	BF283268	283268 B. floridiae	148	47	2.7	989	7	CPCRMT	Y08243 C. pteroides
76	48	2.8	8970	36	BMAYNLRR	U07847 Bombyx mori	149	47	2.7	1566	7	CRO7333	AJ007333 Catharant
77	48	2.8	1248	36	HMRAS1	X70839 H. magnipapli	150	47	2.7	1686	7	MECGF5	X77462 M. esculenta
78	48	2.8	2312	36	TG093689	U93689 foxoplasma	151	47	2.7	1975	7	NTPOLPHOX	Y12501 Nicotiana t
79	48	2.8	1159	37	AF080131	AF080131 Drosophila	152	47	2.7	1577	7	OSCHSLKN	X18181 O. sativa mr
80	47	2.7	931	37	AF054289	AF054289 Macropus	153	47	2.7	1739	7	PC042442	U12442 P. deltoides
81	47	2.7	2022	3	BTDSPO	X95310 B. taurus mr	154	47	2.7	1305	7	PDCLALDHA	Z19568 P. deltoides
82	47	2.7	4719	3	BRNMXOR	X83508 B. taurus mr	155	47	2.7	572	7	PPY16958	PPY16958 Pterium prat
83	47	2.7	3934	3	RABEMMGX	L04504 Oryctolagus	156	47	2.7	554	7	PPY16960	PPY16960 Pterium prat
84	47	2.7	732	3	SE5585	SE5585 refinol-bin	157	47	2.7	1528	7	PSA2155	PSA2155 Pterium prat
85	47	2.7	2307	4	AF043493	AF043493 Gallus ga	158	47	2.7	1022	7	SPPLA45R	X79904 Phalaenopsi
86	47	2.7	3188	4	AF124396	AF124396 Dantio rer	159	47	2.7	2051	7	SOPRKN	X01355 R. nileveus p
87	47	2.7	4230	4	AF124435	AF124435 Dantio rer	160	47	2.7	2647	7	SOPRKN	X01355 R. nileveus p
88	47	2.7	1584	4	CCRHO	U07475 Cyprinus ca	161	47	2.7	2532	7	SOPRKN	X01355 R. nileveus p
89	47	2.7	4387	4	DRE5690	AI005690 Dantio rer	162	47	2.7	2055	7	TAU48227	Z11528 T. tritium ae
90	47	2.7	3348	4	DRP477POU	X07905 D. rerio mrn	163	47	2.7	1061	7	THIGMR	AF002226 Nicotiana
91	47	2.7	1208	4	GSPMADPR	X83397 Gallus sp.	164	47	2.7	759	8	AF002226	AF002226 Nicotiana
92	47	2.7	1584	4	S74449	S74449 rhodopsin l	165	47	2.7	853	8	AF008214	AF008214 Musa acum
93	47	2.7	1845	4	XLGYCE	Z13966 X. laevis mr	166	47	2.7	868	8	AF049923	AF049923 Petunia x
94	47	2.7	4435	4	XLMPLG	Z97214 Xenopus lae	167	47	2.7	625	8	AF049923	AF049923 Petunia x
95	47	2.7	2937	4	XLMPLG	X59813 X. laevis MP	168	47	2.7	1320	8	AF064552	AF064552 Apium gra
96	47	2.7	2470	4	XLMPLG	U62807 Xenopus lae	169	47	2.7	957	8	AF064552	AF064552 Apium gra
97	47	2.7	1239	5	AO8911	AO8911 H. sapiens	170	47	2.7	1425	8	AF071890	AF071890 Pinus ar
98	47	2.7	102	5	AO8911	AO8911 H. sapiens	171	47	2.7	1043	8	AF076252	AF076252 Arabidops
99	47	2.7	132	5	AO8912	AO8912 H. sapiens	172	47	2.7	605	8	AF079185	AF079185 Arabidops
100	47	2.7	843	5	AI2522	AI2522 Ag169 clone	173	47	2.7	1570	8	AF123482	AF123482 Myosphae
101	47	2.7	1141	5	AF57313	AF57313 Sequence 19	174	47	2.7	1809	8	ATPUR2	U79733 Beta procum
102	47	2.7	1165	5	AF57313	AF57313 Sequence 20	175	47	2.7	1450	8	ATU53864	ATU53864 Arabidops
103	47	2.7	1078	5	AF57313	AF57313 Sequence 19	176	47	2.7	1450	8	ATU53864	ATU53864 Arabidops
104	47	2.7	1243	5	AF57313	AF57313 Sequence 20	177	47	2.7	1360	8	BP079733	BP079733 Arabidops
105	47	2.7	1748	5	AF57313	AF57313 Sequence 19	178	47	2.7	616	8	CTAJ5082	CTAJ5082 Arabidops
106	47	2.7	1141	5	AF57313	AF57313 Sequence 19	179	47	2.7	200000	9	SEU58050	SEU58050 Arabidops
107	47	2.7	767	5	AF57313	AF57313 Sequence 19	180	47	2.7	100000	9	AP000081	AP000081 Homo sapi
108	47	2.7	1141	5	AF57313	AF57313 Sequence 19	181	47	2.7	102456	9	AP000081	AP000081 Homo sapi
109	47	2.7	1604	5	AR009628	AR009628 Sequence	182	47	2.7	100375	9	HS193B12	HS193B12 Homo sapi
110	47	2.7	568	5	AR012379	AR012379 Sequence	183	47	2.7	89328	9	HS396C22	HS396C22 Homo sapi
111	47	2.7	1172	5	AR013797	AR013797 Sequence	184	47	2.7	36669	9	HS433G19	HS433G19 Homo sapi
112	47	2.7	458	5	AR022283	AR022283 Sequence	185	47	2.7	6571	9	HS4012755	HS4012755 Homo sapi
113	47	2.7	1794	5	EO2349	EO2349 cDNA sequen	186	47	2.7	4413	9	HSA012755	HSA012755 Homo sapi
114	47	2.7	2051	5	EO3168	EO3168 Promoter of	187	47	2.7	1976	9	HUMAP21B	HUMAP21B Homo sapien
115	47	2.7	916	5	E12747	E12747 Rhodnius pr	188	47	2.7	3380	10	AF002672	AF002672 Homo sapi
116	47	2.7	1868	5	E12806	E12806 cDNA encodi	189	47	2.7	43761	10	HS5782	HS5782 Homo sapi
117	47	2.7	766	5	IO8319	IO8319 Sequence 1	190	47	2.7	2530	10	HS5782	HS5782 Homo sapi
118	47	2.7	2549	5	IO9499	IO9499 Sequence 20	191	47	2.7	1678	10	HS5782	HS5782 Homo sapi
119	47	2.7	1172	5	IO9499	IO9499 Sequence 20	192	47	2.7	1598	10	HS5782	HS5782 Homo sapi
120	47	2.7	1172	5	IO9499	IO9499 Sequence 20	193	47	2.7	1500	10	HS5782	HS5782 Homo sapi
121	47	2.7	399	5	IO9499	IO9499 Sequence 20	194	47	2.7	1606	10	HS5782	HS5782 Homo sapi
122	47	2.7	399	5	IO9499	IO9499 Sequence 20	195	47	2.7	3174	10	HS5782	HS5782 Homo sapi
123	47	2.7	1172	5	IO9499	IO9499 Sequence 20	196	47	2.7	540	10	HS5782	HS5782 Homo sapi
124	47	2.7	1364	5	IO9499	IO9499 Sequence 20	197	47	2.7	1252	10	HS5782	HS5782 Homo sapi
125	47	2.7	1393	5	IO9499	IO9499 Sequence 20	198	47	2.7	2790	10	HS5782	HS5782 Homo sapi
126	47	2.7	3214	5	IO9499	IO9499 Sequence 20	199	47	2.7	71	10	HS5782	HS5782 Homo sapi
127	47	2.7	399	5	IO9499	IO9499 Sequence 20	200	47	2.7	207957	11	AC004470	AC004470 Homo sapi
128	47	2.7	3214	5	IO9499	IO9499 Sequence 20	201	47	2.7	133478	11	AC004470	AC004470 Homo sapi
129	47	2.7	120	5	IO9499	IO9499 Sequence 20	202	47	2.7	167525	11	AC005224	AC005224 Homo sapi
130	47	2.7	966	5	IO9499	IO9499 Sequence 20	203	47	2.7	2745	11	AF060555	AF060555 Homo sapi
131	47	2.7	120	5	IO9499	IO9499 Sequence 20	204	47	2.7	2378	11	AF060555	AF060555 Homo sapi
132	47	2.7	120	5	IO9499	IO9499 Sequence 20	205	47	2.7	1809	11	AF060555	AF060555 Homo sapi
133	47	2.7	1364	5	IO9499	IO9499 Sequence 20	206	47	2.7	4308	11	AF060555	AF060555 Homo sapi
134	47	2.7	144	5	IO9499	IO9499 Sequence 20	207	47	2.7	4459	11	AF060555	AF060555 Homo sapi
135	47	2.7	11500	6	AF031901	AF031901 Bacterioph	208	47	2.7	586	12	AF038562	AF038562 Mus muscu
136	47	2.7	1929	7	AB025714	AB025714 Nicotiana	209	47	2.7	2252	12	AF047716	AF047716 Mus muscu
137	47	2.7	1042	7	AB027501	AB027501 Dancus ca	210	47	2.7	1591	12	AF087943	AF087943 Rattus no
138	47	2.7	640	7	AB027501	AB027501 Dancus ca	211	47	2.7	1181	12	AF087943	AF087943 Rattus no
139	47	2.7	1797	7	AB027501	AB027501 Dancus ca	212	47	2.7	1591	12	AF087943	AF087943 Rattus no
140	47	2.7	440	7	AB027501	AB027501 Dancus ca	213	47	2.7	4130	12	AF139373	AF139373 Mus muscu
141	47	2.7	1744	7	AB027501	AB027501 Dancus ca	214	47	2.7	1180	12	AF139373	AF139373 Mus muscu
142	47	2.7	1694	7	AB027501	AB027501 Dancus ca	215	47	2.7	1206	12	AF139373	AF139373 Mus muscu
143	47	2.7	2245	7	AB027501	AB027501 Dancus ca	216	47	2.7	1584	12	AF139373	AF139373 Mus muscu
144	47	2.7	2386	7	AB027501	AB027501 Dancus ca	217	47	2.7	1180	12	AF139373	AF139373 Mus muscu
145	47	2.7	848	7	AB027501	AB027501 Dancus ca	218	47	2.7	1422	12	AF139373	AF139373 Mus muscu
146	47	2.7	882	7	AB027501	AB027501 Dancus ca	219	47	2.7	3770	12	AF139373	AF139373 Mus muscu
147	47	2.7	1435	7	AB027501	AB027501 Dancus ca	220	47	2.7	1575	12	AF139373	AF139373 Mus muscu

221	47	2.7	2398	12	MMTAFI48	Y09972 M.musculus	294	46	2.6	1558	4	XLBMP4	X63426 X.laevis mr
222	47	2.7	2179	12	MMTFERES	X57796 Mouse mRNA	295	46	2.6	1389	4	XLBNPNO38	X56039 X.laevis mr
223	47	2.7	1858	12	MMTRANS1	X63162 M.musculus	296	46	2.6	2700	4	XLOOCM	X68249 X.laevis oo
224	47	2.7	5330	12	MMU49908	U46908 Mus musculus	297	46	2.6	3905	5	A03736	A03736 H.sapiens m
225	47	2.7	4032	12	MMU79523	U79523 Mus musculus	298	46	2.6	80	5	A08907	A08907 H.sapiens (
226	47	2.7	642	12	MMU95114	U95114 Mus musculus	299	46	2.6	77	5	A08908	A08908 H.sapiens (
227	47	2.7	2621	12	MUSBREDA	L13397 Mus musculus	300	46	2.6	96	5	A08909	A08909 H.sapiens (
228	47	2.7	3924	12	MUSUNKNE	L04849 Mouse (clon	301	46	2.6	104	5	A08910	A08910 H.sapiens (
229	47	2.7	1056	12	MUSUNKNE	L04849 Mouse (clon	302	46	2.6	161	5	A08916	A08916 H.sapiens (
230	47	2.7	1262	12	MUSUNKNE	L04849 Mouse (clon	303	46	2.6	941	5	A15345	A15345 Bam HI huma
231	47	2.7	1395	12	RNCAALIC	X66871 R.norvegicu	304	46	2.6	748	5	A21101	A21101 N.tabacum S
232	47	2.7	3392	12	RNCPHYRN	X60769 Rat sfd mrn	305	46	2.6	347	5	A23327	A23327 O.sativa mr
233	47	2.7	1194	12	RNCPHYRN	S53987 nicotinic r	306	46	2.6	796	5	A23630	A23630 Corn anthr
234	47	2.7	3030	12	RNCPHYRN	S83440 RNH-1-prepr	307	46	2.6	1830	5	A27631	A27631 JGMV coat p
235	47	2.7	2756	12	HSU27131	U27131 Human STS 1	308	46	2.6	1014	5	A30330	A30330 Artificial
236	47	2.7	492	14	HSU27131	U27131 Human STS 1	309	46	2.6	1016	5	A30331	A30331 Artificial
237	47	2.7	566	14	HSU28784	U28784 Human STS 1	310	46	2.6	3905	5	A31790	A31790 H.sapiens m
238	47	2.7	5482	15	AF090435	AF090435 Cloning v	311	46	2.6	1013	5	A32826	A32826 Synthetic c
239	47	2.7	14225	17	LDVGLYPOL	L13398 Lactate deh	312	46	2.6	1015	5	A32827	A32827 Synthetic c
240	47	2.7	961	17	MDU76034	U76034 Microplitis	313	46	2.6	1837	5	A47357	A47357 Sequence 4
241	47	2.7	1801	17	SCU78817	U78817 Saccharomyc	314	46	2.6	1934	5	A49139	A49139 Sequence 4
242	47	2.7	11141	17	TEU27491	U27491 Tick-borne	315	46	2.6	3437	5	A52184	A52184 Sequence 1
243	47	2.7	11141	17	TEU27491	U27491 Tick-borne	316	46	2.6	13414	5	A58184	A58184 Sequence 17
244	47	2.7	152143	34	AC004932	AC004932 Homo sapi	317	46	2.6	10288	5	A58523	A58523 Sequence 27
245	47	2.7	131338	34	HSU773A18	AL049557 Homo sapi	318	46	2.6	2400	5	A58524	A58524 Sequence 28
246	47	2.7	312766	34	PFMALP4	AL035476 plasmodu	319	46	2.6	6791	5	A60092	A60092 Sequence 1
247	47	2.7	182538	34	PFMALP4	AL035477 plasmodu	320	46	2.6	10195	5	A60094	A60094 Sequence 4
248	47	2.7	228625	35	AC004153	AC004153 plasmodu	321	46	2.6	1577	5	A65336	A65336 Sequence 59
249	47	2.7	187544	35	AC004157	AC004157 plasmodu	322	46	2.6	938	5	A70386	A70386 Sequence 16
250	47	2.7	180278	35	AC007445	AC007445 Homo sapi	323	46	2.6	1001	5	AR016469	AR016469 Sequence
251	47	2.7	160021	35	AC007857	AC007857 Homo sapi	324	46	2.6	2267	5	E01314	E01314 CDNA encod1
252	47	2.7	4875	36	AB009080	AB009080 Dictyoste	325	46	2.6	1924	5	E01356	E01356 CDNA encod1
253	47	2.7	14955	36	AF000580	AF000580 Dictyoste	326	46	2.6	1861	5	E01357	E01357 CDNA encod1
254	47	2.7	2253	36	AF032396	AF032396 Hyalophor	327	46	2.6	988	5	E01573	E01573 CDNA sequen
255	47	2.7	1743	36	CELRABGDI	U00002 Guanine nuc	328	46	2.6	988	5	E01575	E01575 CDNA sequen
256	47	2.7	10053	36	CFU19151	U19151 Crithidia f	329	46	2.6	3851	5	E01614	E01614 DNA sequen
257	47	2.7	678	36	DE5468	X53445 Dictyostei1	330	46	2.6	3820	5	E02221	E02221 CDNA encod1
258	47	2.7	681	36	DDE628	X53448 Dictyostei1	331	46	2.6	958	5	E02319	E02319 DNA encod1
259	47	2.7	1840	36	DDU73686	U73686 Dictyostei1	332	46	2.6	1558	5	E03671	E03671 DNA encod1
260	47	2.7	3937	36	DMU35074	U35074 Drosophila	333	46	2.6	667	5	E06788	E06788 DNA encod1
261	47	2.7	1432	36	PAT69957	U69957 Periplaneta	334	46	2.6	688	5	E06789	E06789 CDNA encod1
262	47	2.7	108908	36	PFMAL3P8	AL034560 Plasmodu	335	46	2.6	667	5	E06790	E06790 Modifie'd DN
263	47	2.7	915	36	RPI61144	U61144 Rhodnius pr	336	46	2.6	9589	5	E07362	E07362 CDNA of Hep
264	47	2.7	1963	36	SMU86674	U61144 Rhodnius pr	337	46	2.6	826	5	E08516	E08516 DNA encod1
265	47	2.7	1288	36	TBBDPMSYN	254162 T.brucei br	338	46	2.6	779	5	E08517	E08517 DNA encod1
266	47	2.7	587	36	TOUT89944	U89944 Taenia ovis	339	46	2.6	2430	5	E12585	E12585 CDNA encod1
267	47	2.7	3351	37	AF031517	AF031517 Drosophill	340	46	2.6	826	5	E12587	E12587 CDNA encod1
268	47	2.7	1104	37	AF047611	AF047611 Euroglyph	341	46	2.6	779	5	E12888	E12888 CDNA encod1
269	47	2.7	1708	37	AF060449	AF060449 Limaex ita	342	46	2.6	1113	5	E13364	E13364 Plasmid pyg
270	47	2.7	5117	37	AF067198	AF067198 Dictyoste	343	46	2.6	1769	5	E13524	E13524 3' untransla
271	47	2.7	1412	37	AF083228	AF083228 Caenorhab	344	46	2.6	622	5	E13525	E13525 3' untransla
272	47	2.7	1509	37	AFI10676	AFI10676 Dictyoste	345	46	2.6	3851	5	E13540	E13540 Seriola sp.
273	47	2.7	3488	37	AFI129433	AFI129433 Drosophill	346	46	2.6	3451	5	E13998	E13998 Cyamidum c
274	47	2.7	855	37	AFI30840	AFI30840 Rhyzopert	347	46	2.6	941	5	E10396	E10396 Sequence 8
275	47	2.7	2856	37	AFI32149	AFI32149 Drosophill	348	46	2.6	688	5	E10396	E10396 Sequence 8
276	47	2.7	928	37	AFI32151	AFI32151 Drosophill	349	46	2.6	1124	5	E10696	E10696 Sequence 11
277	47	2.7	1852	37	AFI32556	AFI32556 Drosophill	350	46	2.6	2035	5	E108608	E108608 Sequence 3
278	47	2.7	3143	37	S73271	S73271 trehalase l	351	46	2.6	1923	5	E109202	E109202 Sequence 1
279	47	2.7	1006	37	TVH87097	U87097 Trichomonas	352	46	2.6	1423	5	E109360	E109360 Sequence 1
280	47	2.7	107402	42	AC007043	AC007043 Homo sapi	353	46	2.6	1013	5	E11571	E11571 Sequence 23
281	47	2.7	2821	3	AF039137	AF039137 Felis cat	354	46	2.6	9589	5	E11861	E11861 Sequence 2
282	47	2.6	2815	3	AF039138	AF039138 Felis cat	355	46	2.6	570	5	E11835	E11835 Sequence 10
283	47	2.6	2405	3	B1LACTRA	X57084 Bovine mRNA	356	46	2.6	570	5	E12142	E12142 Sequence 10
284	47	2.6	2714	3	BITU7504	U75304 Bos laurus	357	46	2.6	2760	5	E125048	E125048 Sequence 30
285	47	2.6	1316	3	OCU131955	AJ131955 Oryctolag	358	46	2.6	1013	5	E125048	E125048 Sequence 30
286	47	2.6	2836	3	SSU49434	U49434 Sus scrofa	359	46	2.6	796	5	E133111	E133111 Sequence 3
287	47	2.6	1665	4	AF080068	AF080068 Xenopus 1	360	46	2.6	1441	5	E13392	E13392 Sequence 8
288	47	2.6	2085	4	AF080062	AF080062 Dario rer	361	46	2.6	609	5	E13492	E13492 Sequence 2
289	47	2.6	1020	4	DRHPRABIN	X97332 D. rerio mrn	362	46	2.6	570	5	E140331	E140331 Sequence 10
290	47	2.6	1714	4	GDSSDBF	X80340 G.gomeseticu	363	46	2.6	1817	5	I42402	I42402 Sequence 5
291	47	2.6	1455	4	MSUY18860	Y18860 Mullus surm	364	46	2.6				
292	47	2.6	1559	4	OLJ000937	AJ000937 Oryzias l	365	46	2.6				
293	47	2.6	2262	4	XELCDC25AA	M96857 Xenopus lae	366	46	2.6				

357	46	2.5	748	5	I44072	144072	Sequence 3	C	440	46	2.6	107135	9	HS145B12	AL008706 Human DNA s
358	46	2.6	1898	5	I46765	146765	Sequence 2	C	441	46	2.6	36429	9	HS14B7	249258 Human DNA s
359	46	2.6	147730	5	I47730	147730	Sequence 1	C	442	46	2.6	136144	9	HS154K9	294277 Human DNA s
360	46	2.6	148978	5	I48978	148978	Sequence 5	C	443	46	2.6	103146	9	HS187N21	298036 Human DNA s
361	46	2.6	240	5	I48979	148979	Sequence 6	C	444	46	2.6	120376	9	HS271G9	292543 Human DNA s
362	46	2.6	289	5	I63077	163077	Sequence 3	C	445	46	2.6	143747	9	HS370M22	282206 Human DNA s
363	46	2.6	1558	5	I66401	166401	Sequence 7	C	446	46	2.6	170952	9	HS501N12	AL022170 Homo sapi
364	46	2.6	5173	5	I68732	168732	Sequence 1	C	447	46	2.6	176433	9	HS711L6	AL022165 Human DNA
365	46	2.6	2405	5	I73428	173428	Sequence 30	C	448	46	2.6	3723	9	HSADTG	Y12226 H. sapiens m
366	46	2.6	216	5	I80062	180062	Sequence 34	C	449	46	2.6	2000	9	HSX28ORF	X99270 H. sapiens x
367	46	2.6	69142	7	AB012239	AB012239	Arabidops	C	450	46	2.6	4782	9	HXM130LEU	M92439 Human Leuci
368	46	2.6	2440	7	AB021862	AB021862	Cucumis m	C	451	46	2.6	2287	9	HUMAZPIA	D00174 Homo sapien
369	46	2.6	1206	7	ATANNEX	X99224 A. thaliana	C	452	46	2.6	2271	9	HUMASP	J05032 Human aspar	
370	46	2.6	513	7	ATCOR66M	X55053 A. thaliana	C	453	46	2.6	1266	9	HUMDP2M	L40386 Human DP-2	
371	46	2.6	835	7	ATP27A	Y11792 A. thaliana	C	454	46	2.6	1659	9	HUMP57B	D44497 Human trans	
372	46	2.6	680	7	ATSRP14	Y10116 A. thaliana	C	455	46	2.6	6075	9	HUMTNL12	M80340 Human trans	
373	46	2.6	1770	7	BNPML2	X73849 B. napus (PN	C	456	46	2.6	2549	9	U00686	U00686 Human putat	
374	46	2.6	1104	7	BOM1PATCP	X93639 B. oleracea	C	457	46	2.6	134218	10	HS170F5	AL035067 Human DNA	
375	46	2.6	1419	7	CCCHSMR	Z67988 C. chinensis	C	458	46	2.6	117338	10	HS173D1	AL031984 Human DNA	
376	46	2.6	783	7	CCFCPI0	AJ000972 Cyclostell	C	459	46	2.6	69017	10	HS410I8	AL0341732 Human DNA	
377	46	2.6	987	7	CLJACOIV	X57832 C. lacryma-j	C	460	46	2.6	101270	10	HS483K16	AL034374 Human DNA	
378	46	2.6	673	7	CRPS27	X83694 C. reinhardt	C	461	46	2.6	130467	10	HS798A17	AL031274 Human DNA	
379	46	2.6	1839	7	DI4059	DI4059 Ananas como	C	462	46	2.6	1608	10	HSDAP3	X83544 H. sapiens D	
380	46	2.6	1735	7	DCRCBT3	Z84571 D. carayophyll	C	463	46	2.6	39015	10	HSE90C2	282182 Homo sapien	
381	46	2.6	2263	7	HVRNAGAM1	X87690 H. vulgare m	C	464	46	2.6	5959	10	HSB8M	X72889 H. sapiens h	
382	46	2.6	2641	7	LENLEPRIB	D14489 Lentinus ed	C	465	46	2.6	1817	10	HSICAM3RN	X69819 H. sapiens I	
383	46	2.6	1236	7	LOLOPIB	M59163 Lolium perit	C	466	46	2.6	5994	10	HSINTB4R	X53587 Human mRNA	
384	46	2.6	1810	7	MCPTKXINA	Z30329 M. crystalli	C	467	46	2.6	3688	10	HSKCHMR	X65873 H. sapiens m	
385	46	2.6	1162	7	MNRBCSA	M95930 Manihot esc	C	468	46	2.6	1483	10	HSMB800068	AL049283 Homo sapi	
386	46	2.6	837	7	NPSODM	X55974 N. plumbagin	C	469	46	2.6	1372	10	HSMB800085	AL049300 Homo sapi	
387	46	2.6	2400	7	NPZEXANT	X95732 N. plumbagin	C	470	46	2.6	580	10	HSMB800144	AL050277 Homo sapi	
388	46	2.6	2851	7	OSGOG1	Y12594 O. sativa mr	C	471	46	2.6	3386	10	HSMB800161	AL050092 Homo sapi	
389	46	2.6	4120	7	OSGOG2	Y12595 O. sativa mr	C	472	46	2.6	1694	10	HSMB800167	AL050381 Homo sapi	
390	46	2.6	875	7	OSRTA2PR	AJ001265 Oryza sat	C	473	46	2.6	669	10	HSMB800251	AL049466 Homo sapi	
391	46	2.6	1391	7	PCU10306	U10306 Phanerocha	C	474	46	2.6	1438	10	HSMB800279	AL049466 Homo sapi	
392	46	2.6	2159	7	PSDPS	Y08611 P. sativum m	C	475	46	2.6	2785	10	HSMB800418	AL050108 Homo sapi	
393	46	2.6	575	7	PSGAMRA	X65154 P. sativum G	C	476	46	2.6	2462	10	HSMB800419	AL050116 Homo sapi	
394	46	2.6	708	7	PSGDCPG	X54377 P. sativum g	C	477	46	2.6	3469	10	HSMB800438	AL050413 Homo sapi	
395	46	2.6	518	7	QSAJ692	AJ000692 Quercus s	C	478	46	2.6	2765	10	HSMB800443	AL050174 Homo sapi	
396	46	2.6	498	7	S47413	S47413 glycine-ric	C	479	46	2.6	959	10	HSMPZE	X61970 H. sapiens m	
397	46	2.6	900	7	SLDEHYDLA	Z21500 S. longipes	C	480	46	2.6	380	10	HSMUC5BR	Y10080 H. sapiens M	
398	46	2.6	2235	7	SOPOMRN	Z66559 S. oleracea	C	481	46	2.6	2300	10	HSMXR7MR	Z37987 H. sapiens m	
399	46	2.6	3437	7	SOPULSPO	X83969 S. oleracea	C	482	46	2.6	1550	10	HSMXR9	Z49216 H. sapiens m	
400	46	2.6	795	7	SOYSCAM1X	L01430 Soybean cal	C	483	46	2.6	39000	10	HSN6D4	Z82250 Human DNA s	
401	46	2.6	1432	7	SPCIG2	X70046 S. pombe mrn	C	484	46	2.6	2739	10	HSTRANSLI	X78627 H. sapiens m	
402	46	2.6	2062	7	WHTEPI1X	M90077 Wheat trans	C	485	46	2.6	1368	10	HSY16645	Y16645 Homo sapien	
403	46	2.6	1016	8	AF005158	AF005158 Arabidops	C	486	46	2.6	2123	10	S38729	S38729 Ku autoanti	
404	46	2.6	610	8	AF009959	AF009959 Oryza sat	C	487	46	2.6	107	10	S77771S1	S77771 aspartylgltu	
405	46	2.6	1592	8	AF019630	AF019630 Magnapor	C	488	46	2.6	837	10	S78214	S78214 APC-tumor s	
406	46	2.6	764	8	AF022731	AF022731 Oryza sat	C	489	46	2.6	125590	11	AC002382	AC002382 Human BAC	
407	46	2.6	1032	8	AF043538	AF043538 Arabidops	C	490	46	2.6	128915	11	AC002416	AC002416 Human Chr	
408	46	2.6	747	8	AF055372	AF055372 Arabidops	C	491	46	2.6	123805	11	AC002464	AC002464 Human BAC	
409	46	2.6	1663	8	AF058763	AF058763 zea may	C	492	46	2.6	41084	11	AC002500	AC002500 Human Cos	
410	46	2.6	2251	8	AF060518	AF060518 Cuphea pu	C	493	46	2.6	197900	11	AC002531	AC002531 Homo sapi	
411	46	2.6	2323	8	AF071888	AF071888 Prunus ar	C	494	46	2.6	70851	11	AC002540	AC002540 Human BAC	
412	46	2.6	1700	8	AF121261	AF121261 Lilium lo	C	495	46	2.6	94320	11	AC002564	AC002564 Human BAC	
413	46	2.6	1647	8	AF124842	AF124842 Capsicum	C	496	46	2.6	45084	11	AC003005	AC003005 Human DNA	
414	46	2.6	1404	8	ATU62741	U62741 Arabidops	C	497	46	2.6	102818	11	AC003042	AC003042 Homo sapi	
415	46	2.6	1492	8	HAGU91341	U91341 Helianthus	C	498	46	2.6	113457	11	AC003686	AC003686 Homo sapi	
416	46	2.6	801	8	RICLMMHSP	M80186 Oryza sativ	C	499	46	2.6	154252	11	AC004062	AC004062 Homo sapi	
417	46	2.6	1276	8	S59422	S59422 Populus x c	C	500	46	2.6	42007	11	AC004200	AC004200 Homo sapi	
418	46	2.6	826	8	S65048	S65048 self-incomp	C	501	46	2.6	156461	11	AC004227	AC004227 Homo sapi	
419	46	2.6	1425	8	SOYSPK3	L19361 Glycine max	C	502	46	2.6	185516	11	AC004383	AC004383 Human Chr	
420	46	2.6	1154	8	TAU87163	U87163 Triticum m	C	503	46	2.6	185516	11	AC004485	AC004485 Homo sapi	
421	46	2.6	3901	8	VUA225088	AJ225088 Vigna ung	C	504	46	2.6	80659	11	AC004553	AC004553 Homo sapi	
422	46	2.6	45966	9	AP000028	AP000028 Homo sapi	C	505	46	2.6	127824	11	AC004822	AC004822 Homo sapi	
423	46	2.6	100000	9	AP000130	AP000130 Homo sapi	C	506	46	2.6	85058	11	AC004989	AC004989 Homo sapi	
424	46	2.6	100000	9	AP000208	AP000208 Homo sapi	C	507	46	2.6	73899	11	AC005156	AC005156 Homo sapi	
425	46	2.6	9918	9	AP000218	AP000218 Homo sapi	C	508	46	2.6	198582	11	AC005291	AC005291 Homo sapi	
426	46	2.6	115118	9	AP000247	AP000247 Homo sapi	C	509	46	2.6	131070	11	AC005295	AC005295 Homo sapi	
427	46	2.6	35932	9	AP000340	AP000340 Homo sapi	C	510	46	2.6	87114	11	AC005353	AC005353 Homo sapi	
428	46	2.6	1681	9	D89079	D89079 Homo sapien	C	511	46	2.6	16167	11	AC005483	AC005483 Homo sapi	
429	46	2.6	1216	9	EN1838	AJ001838 Homo sapi	C	512	46	2.6	183085	11	AC005815	AC005815 complet	

513	2.6	134045	11	AC005886	AC005886 Homo sapi	586	2.6	417	17	TBEV3UT	Y08865 Tick-borne
514	2.6	178367	11	AC006115	AC006115 Homo sapi	587	2.6	768	17	TEU72490	Y27490 Tick-borne
515	2.6	154354	11	AC006197	AC006197 Homo sapi	588	2.6	9589	26	E07361	E07361 gRNA of Hep
516	2.6	79085	11	AC006221	AC006221 Homo sapi	589	2.6	1657	26	E11332	E11332 Human cDNA
517	2.6	220218	11	AC006288	AC006288 Homo sapi	590	2.6	32208	34	AC004105	AC004105 Homo sapi
518	2.6	186555	11	AC006313	AC006313 Homo sapi	591	2.6	206056	34	AC004387	AC004387 Homo sapi
519	2.6	1728	11	AF013249	AF013249 Homo sapi	592	2.6	170000	34	AC004579	AC004579 Homo sapi
520	2.6	2150	11	AF017790	AF017790 Homo sapi	593	2.6	190000	34	AC004560	AC004560 Homo sapi
521	2.6	1572	11	AF035305	AF035305 Homo sapi	594	2.6	198477	34	AC004971	AC004971 Homo sapi
522	2.6	2632	11	AF040751	AF040751 Homo sapi	595	2.6	171188	34	AC005025	AC005025 Homo sapi
523	2.6	2843	11	AF044323	AF044323 Homo sapi	596	2.6	180593	34	AC005282	AC005282 Homo sapi
524	2.6	2098	11	AF060866	AF060866 Homo sapi	597	2.6	146209	34	AC005522	AC005522 Homo sapi
525	2.6	2619	11	AF068229	AF068229 Homo sapi	598	2.6	41755	34	AC006179	AC006179 Homo sapi
526	2.6	5075	11	AF081195	AF081195 Homo sapi	599	2.6	195764	34	AC006483	AC006483 Homo sapi
527	2.6	2802	11	AF084644	AF084644 Homo sapi	600	2.6	124673	34	HS29H4	AL078475 Homo sapi
528	2.6	2905	11	AF084645	AF084645 Homo sapi	601	2.6	112555	34	HS419C19	AL035407 Homo sapi
529	2.6	1212	11	AF100781	AF100781 Homo sapi	602	2.6	115272	34	HS475N16	AL035587 Homo sapi
530	2.6	5827	11	HSMH345	HSMH345 Homo sapien	603	2.6	91029	34	HS72E17	AL033523 Homo sapi
531	2.6	2902	11	HSU31501	HSU31501 Human fragi	604	2.6	133505	34	HS785G19	AL035458 Homo sapi
532	2.6	2516	11	HSU37359	HSU37359 Homo sapien	605	2.6	169254	34	HS892F13	AL009183 Homo sapi
533	2.6	5173	11	HSU38847	HSU38847 Homo TAR R	606	2.6	142807	34	HSJ02553	AL002553 Homo sapi
534	2.6	1280	11	HSU54559	HSU54559 Homo sapien	607	2.6	110000	34	HSJ068N3-1	Continuation (2 of
535	2.6	2814	11	HSU73682	HSU73682 Human menin	608	2.6	263628	34	HSJ0784K2	AL078588 Homo sapi
536	2.6	42179	11	HSU76377	HSU76377 Human Olfac	609	2.6	247159	34	HSJ081F6	AL049762 Homo sapi
537	2.6	8745	11	HSU88966	HSU88966 Human prote	610	2.6	123100	34	HSJ727110	AL050332 Homo sapi
538	2.6	4082	11	HSU91329	HSU91329 Human kines	611	2.6	111165	34	HSJ842K24	AL050310 Homo sapi
539	2.6	69015	11	HSU95739	HSU95739 Human chrom	612	2.6	300172	35	AC005308	AC005308 Plasmodi
540	2.6	188636	11	HUAC002287	HUAC002287 Homo sapi	613	2.6	104992	35	AC005504	AC005504 Plasmodi
541	2.6	130491	11	HUAC003977	HUAC003977 Homo sapi	614	2.6	288033	35	AC005507	AC005507 Plasmodi
542	2.6	38588	12	AC004389	AC004389 Mus muscu	615	2.6	157771	35	AC005902	AC005902 Homo sapi
543	2.6	4128	12	AF002263	AF002263 Mus muscu	616	2.6	169387	35	AC006338	AC006338 Homo sapi
544	2.6	2246	12	AF026124	AF026124 Mus muscu	617	2.6	205380	35	AC006373	AC006373 Homo sapi
545	2.6	10919	12	AF031572	AF031572 Mus muscu	618	2.6	195600	35	AC006460	AC006460 Homo sapi
546	2.6	4452	12	AF054831	AF054831 Mus muscu	619	2.6	122459	35	AC006944	AC006944 Mus muscu
547	2.6	3895	12	AF079765	AF079765 Mus muscu	620	2.6	122459	35	AC006944	AC006944 Mus muscu
548	2.6	1661	12	AF081366	AF081366 Rattus no	621	2.6	106633	35	AC006994	AC006994 Mus muscu
549	2.6	1315	12	AF082526	AF082526 Mus muscu	622	2.6	195550	35	AC006995	AC006995 Homo sapi
550	2.6	965	12	AF106945	AF106945 Rattus no	623	2.6	171771	35	AC007038	AC007038 Homo sapi
551	2.6	4365	12	AF114168	AF114168 Mus muscu	624	2.6	182480	35	AC007056	AC007056 Homo sapi
552	2.6	834	12	AF115410	AF115410 Cricetulu	625	2.6	168457	35	AC007078	AC007078 Homo sapi
553	2.6	1855	12	AF143957	AF143957 Mus muscu	626	2.6	183800	35	AC007391	AC007391 Homo sapi
554	2.6	154614	12	AF153957	AF153957 Mus muscu	627	2.6	193301	35	AC007392	AC007392 Homo sapi
555	2.6	251364	12	MMAC000664	MMAC000664 Mus muscu	628	2.6	187817	35	AC007597	AC007597 Homo sapi
556	2.6	158405	12	MMHC310M6	MMHC310M6 Mus muscu	629	2.6	167373	35	AC007604	AC007604 Homo sapi
557	2.6	2997	12	MMPC1A3A	MMPC1A3A Mus muscu	630	2.6	328876	35	AC007610	AC007610 Homo sapi
558	2.6	1778	12	MMU02885	MMU02885 Mus muscu	631	2.6	191300	35	AC007749	AC007749 Homo sapi
559	2.6	3715	12	MMU35846	MMU35846 Mus muscu	632	2.6	925	35	AEV0440X	AEV0440X Homo sapi
560	2.6	2693	12	MMU83172	MMU83172 Mus muscu	633	2.6	4028	36	AF003106	AF003106 Drosophila
561	2.6	756	12	MRBR17B	MRBR17B Mus muscu	634	2.6	2062	36	AF035482	AF035482 Trichoptu
562	2.6	901	12	MUSMASC3H2	MUSMASC3H2 Mouse mast	635	2.6	40753	36	BFY18367	BFY18367 Branchiosto
563	2.6	795	12	MUSUNKNL	MUSUNKNL Mouse	636	2.6	1355	36	CEU68266	CEU68266 Caenorhabd
564	2.6	814	12	RATPHGP	RATPHGP Rattus norv	637	2.6	763	36	CEU96172	CEU96172 C. viridis m
565	2.6	1461	12	RNMAFA	RNMAFA R. norvegicu	638	2.6	1907	36	DCV0X6R	DCV0X6R Homo sapi
566	2.6	409	12	RNMTCYTOX	RNMTCYTOX R. norvegicu	639	2.6	479	36	DDRE	DDRE Homo sapi
567	2.6	1837	12	RNP2XMR	RNP2XMR R. norvegicu	640	2.6	6428	36	DDIACIN8A	DDIACIN8A Homo sapi
568	2.6	2133	12	RNP4502C	RNP4502C Rat mRNA to	641	2.6	91	36	DDIGP138A	DDIGP138A Homo sapi
569	2.6	2325	12	RNP03290	RNP03290 Rattus norv	642	2.6	2581	36	DDIPORIN	DDIPORIN Homo sapi
570	2.6	1560	12	RNU76419	RNU76419 Rattus norv	643	2.6	1128	36	DIRRSS	DIRRSS Homo sapi
571	2.6	1837	12	S69385	S69385 epithelial	644	2.6	1946	36	DMRNL	DMRNL Homo sapi
572	2.6	569407	12	S79832	S79832 Edg+endogli	645	2.6	2415	36	DMR55MR	DMR55MR Homo sapi
573	2.6	3180	12	S83436	S83436 GTP-binding	646	2.6	5830	36	DMSCPRTH	DMSCPRTH Homo sapi
574	2.6	1165	12	S83436	S83436 GTP-binding	647	2.6	3504	36	DMRNL	DMRNL Homo sapi
575	2.6	371	14	AMU73923	AMU73923 SHGC-54582	648	2.6	140	36	DMRNL	DMRNL Homo sapi
576	2.6	432	14	G36768	G36768 SHGC-54582	649	2.6	1428	36	DMR55MR	DMR55MR Homo sapi
577	2.6	8392	15	CVU30497	CVU30497 Cloning vec	650	2.6	1376	36	DMSCPRTH	DMSCPRTH Homo sapi
578	2.6	8392	15	CVU30497	CVU30497 Cloning vec	651	2.6	1482	36	HERFAMIDE	HERFAMIDE Homo sapi
579	2.6	791	16	OSA004866	OSA004866 Oryza sat	652	2.6	631	36	HVU48594	HVU48594 Homo sapi
580	2.6	1208	17	HPCHCUST70	HPCHCUST70 Hepatitis C	653	2.6	1601	36	LPACT	LPACT Homo sapi
581	2.6	1227	17	HPCHCUST70	HPCHCUST70 Hepatitis C	654	2.6	1567	36	LPACT	LPACT Homo sapi
582	2.6	9589	17	HPCHCUST70	HPCHCUST70 Hepatitis C	655	2.6	1577	36	LPACT	LPACT Homo sapi
583	2.6	9456	17	HPCHCUST70	HPCHCUST70 Hepatitis C	656	2.6	1577	36	LPACT	LPACT Homo sapi
584	2.6	9456	17	HPCHCUST70	HPCHCUST70 Hepatitis C	657	2.6	1577	36	LPACT	LPACT Homo sapi
585	2.6	12226	17	RTU0LP	RTU0LP Rice tungro	658	2.6	2633	36	PFA5PHM	PFA5PHM Homo sapi

659	46	2.6	3528	36	PF016995	U16995 Plasmodium	732	45	2.6	2123	4	GGU77351	U77351 Gallus gall
C 660	46	2.6	1538	36	PF097128	U97128 Plasmodium	733	45	2.6	1620	4	GGVITRO	Y11030 G.gallus mr
C 661	46	2.6	476	36	PFY18125	Y18125 Physarum po	734	45	2.6	1170	4	HNPROLAC3	X61040 H.nobilis m
662	46	2.6	869	36	S75286	S75286 major aller	735	45	2.6	1544	4	OLANEXM3	Y11254 O.laticpes m
663	46	2.6	613	36	S76337	S76337 IgE-binding	736	45	2.6	2728	4	OMAJ3118	AY003118 Oncorhyn
664	46	2.6	1059	36	S76894	S76894 Daep-1-cuti	737	45	2.6	586	4	OMR33980	U83980 Oncorhynch
C 665	46	2.6	2532	36	SLMACTARD	M15272 slime mold	738	45	2.6	443	4	PBDESRPB	X72387 P.bicolor m
666	46	2.6	1407	36	SPMUSACT	X61042 S.plicata m	739	45	2.6	958	4	S54890	S54890 growth horm
667	46	2.6	737	36	TCR2A	Y13395 Trypanosoma	740	45	2.6	994	4	S56212	S56212 nerve growt
668	46	2.6	7291	36	TBRPTE	M62862 Trypanosoma	741	45	2.6	1334	4	S59519	S59519 stannocalc
669	46	2.6	953	36	TYPC4	X77221 T.vaginalis	742	45	2.6	5078	4	TNU75604	U75604 Tilapia nil
C 670	46	2.6	13674	36	U00691	U00691 Dictyostei	743	45	2.6	542	4	XEHIS3LA	XEHIS3LA xenopus lae
C 671	46	2.6	13729	36	U00796	U00796 Dictyostei	744	45	2.6	2526	4	XHCFIN	XHCFIN x.laevis B-
C 672	46	2.6	12530	37	AE001417	AE001417 Plasmodiu	745	45	2.6	1469	4	XLB506AP4	XLB506 X.laevis B-
673	46	2.6	2435	37	AF007768	AF007768 Choriston	746	45	2.6	2134	4	XLICCTG	XLICCTG x.laevis Cc
674	46	2.6	790	37	AF040384	AF040384 Schistos	747	45	2.6	1500	4	XLNKATP	Y11587 X.laevis mr
675	46	2.6	845	37	AF047612	AF047612 Euroglyph	748	45	2.6	2729	4	XLU68387	U68387 Xenopus lae
676	46	2.6	1429	37	AF072219	AF072219 Blatella	749	45	2.6	1417	4	ZOPY18678	ZOPY18678 zosterisess
677	46	2.6	2304	37	AF080675	AF080675 Dictyoste	750	45	2.6	1394	5	A07588	A07588 Synthetic H
678	46	2.6	905	37	AF081800	AF081800 Dictyoste	751	45	2.6	2399	5	A07647	A07647 Synthetic g
679	46	2.6	1412	37	AF083227	AF083227 Caenorhab	752	45	2.6	3347	5	A12297	A12297 RII beta cd
680	46	2.6	477	37	AF098065	AF098065 Trypanoso	753	45	2.6	7859	5	A17115	A17115 yeast expe
681	46	2.6	6396	37	AF102268	AF102268 Pacifasta	754	45	2.6	7984	5	A18079	A18079 yeast expe
682	46	2.6	3754	37	AF132141	AF132141 Drosophil	755	45	2.6	80	5	A18777	A18777 5' terminus
683	46	2.6	2330	37	AF132142	AF132142 Drosophil	756	45	2.6	11517	5	A18788	A18788 complete nu
684	46	2.6	2008	37	AF145689	AF145689 Drosophil	757	45	2.6	7859	5	A18996	A18996 SEQ ID NO:
685	46	2.6	446	37	AF146743	AF146743 Mesobuthu	758	45	2.6	1284	5	A20553	A20553 NF-YB. 6/19
686	46	2.6	1636	37	CEU60058	U60058 Caenorhabd	759	45	2.6	1046	5	A21103	A21103 N.cabacum S
687	46	2.6	1459	37	GIU92647	U92647 Geocarcinus	760	45	2.6	3138	5	A21625	A21625 expression
688	46	2.6	1626	37	GIU81605	U81605 Onchocerca	761	45	2.6	2966	5	A27171	A27171 Tyrosine Ki
689	46	2.6	2846	37	TOXATPC	L39079 Toxoplasma	762	45	2.6	1966	5	A31001	A31001 G.gallus DN
C 690	46	2.6	40933	42	AC002471	AC002471 Homo sapi	763	45	2.6	1395	5	A41575	A41575 Sequence 1
691	46	2.6	41958	42	AC005374	AC005374 Homo sapi	764	45	2.6	2594	5	A45787	A45787 Sequence 25
692	46	2.6	176257	42	AC006039	AC006039 Homo sapi	765	45	2.6	2643	5	A47363	A47363 Sequence 10
693	46	2.6	191145	42	AC006112	AC006112 Homo sapi	766	45	2.6	141	5	A51774	A51774 Sequence 9
C 694	46	2.6	126732	42	AC006336	AC006336 Homo sapi	767	45	2.6	3200	5	A52563	A52563 Sequence 4
C 695	46	2.6	80858	42	AC007360	AC007360 Homo sapi	768	45	2.6	1276	5	A57389	A57389 Sequence 1
696	46	2.6	3016	42	AF104032	AF104032 Homo sapi	769	45	2.6	1713	5	A59344	A59344 Sequence 1
C 697	46	2.6	4670	42	AF104032	AF104032 Homo sapi	770	45	2.6	1441	5	A59403	A59403 Sequence 63
698	46	2.6	2180	42	AF106697	AF106697 Homo sapi	771	45	2.6	3114	5	A65340	A65340 Sequence 3
699	46	2.6	1853	42	HSU58996	U58996 Homo sapien	772	45	2.6	1040	5	A65965	A65965 Sequence 25
700	46	2.6	3227	3	AF010191	AF010191 Felis cat	773	45	2.6	1713	5	A65934	A65934 Sequence 2
701	46	2.6	4211	3	AF019298	AF019298 Bos tauru	774	45	2.6	1602	5	AR000496	AR000496 Sequence
702	45	2.6	2494	3	AF026008	AF026008 Oryctolag	775	45	2.6	11517	5	AR001293	AR001293 Sequence
703	45	2.6	3345	3	AF126247	AF126247 Canis fam	776	45	2.6	80	5	AR001312	AR001312 Sequence
704	45	2.6	4697	3	BTTTHOM2	X96540 B.taurus mr	777	45	2.6	3040	5	AR005011	AR005011 Sequence
705	45	2.6	1024	3	BTU79414	U79414 Bos taurus	778	45	2.6	1425	5	AR005195	AR005195 Sequence
706	45	2.6	5425	3	CFP180RC	X87224 Canis famli	779	45	2.6	6671	5	AR011880	AR011880 Sequence
707	45	2.6	3334	3	CFSRP72	X67813 C.familiari	780	45	2.6	1181	5	AR016802	AR016802 Sequence
708	45	2.6	1065	3	CHAF000167	AF000167 Capra hir	781	45	2.6	1147	5	AR020828	AR020828 Sequence
709	45	2.6	442	3	OCRPYC	U58653 Oryctolagus	782	45	2.6	2696	5	AR020905	AR020905 Sequence
710	45	2.6	3485	3	OCU58653	M84133 Sus scrofa	783	45	2.6	2696	5	E00617	E00617 DNA encodi
711	45	2.6	462	3	PICAC3A	M22640 Rabbit MHC	784	45	2.6	2696	5	E00717	E00717 cdna encodi
712	45	2.6	2763	3	RABMHDP3A	X59414 Porcine Apo	785	45	2.6	1235	5	E00778	E00778 Rat liver c
713	45	2.6	693	3	SSAPOAI	AJ131826 Sus scrof	786	45	2.6	669	5	E00984	E00984 Probe dete
714	45	2.6	1573	3	SSC131826	X56503 Porcine CKI	787	45	2.6	1973	5	E01812	E01812 cdna encodi
715	45	2.6	776	3	SSCKTIB	U90884 Sus scrofa	788	45	2.6	960	5	E01963	E01963 DNA encodi
716	45	2.6	2588	3	SSU37312	AF029404 Xenopus 1	789	45	2.6	1014	5	E02152	E02152 DNA encodi
717	45	2.6	1744	3	SSU90884	AF053120 Xenopus 1	790	45	2.6	3630	5	E02152	E02152 human 'A' my
718	45	2.6	1339	3	AF029404	AF059612 Xenopus 1	791	45	2.6	1302	5	E02756	E02756 DNA encodi
719	45	2.6	2161	4	AF053120	AF061981 Xenopus 1	792	45	2.6	1014	5	E02914	E02914 cdna encodi
720	45	2.6	1250	4	AF061981	AF061981 Xenopus 1	793	45	2.6	3429	5	E03202	E03202 DNA encodi
721	45	2.6	2896	4	AF061981	AF061981 Xenopus 1	794	45	2.6	1588	5	E03348	E03348 cdna sequen
722	45	2.6	1554	4	AF061981	AF090432 Danio rer	795	45	2.6	1581	5	E03349	E03349 cdna sequen
723	45	2.6	5138	4	AF090432	AF100931 Oncorhyn	796	45	2.6	1581	5	E04233	E04233 DNA encodi
724	45	2.6	305	4	BJPBP2	X56530 B.japonicus	797	45	2.6	1662	5	E04257	E04257 DNA encodi
725	45	2.6	1770	4	CZPZP2A	M27491 C.carpio mr	798	45	2.6	2339	5	E05822	E05822 DNA encodi
726	45	2.6	1004	4	CHKGR78	M27491 Chicken 78-	799	45	2.6	1632	5	E07108	E07108 cdna encodi
727	45	2.6	4577	4	DRAP1PAT	U89295 Danio rerio	800	45	2.6	9502	5	E07941	E07941 CDNA of Hep
728	45	2.6	1111	4	GNKX28	Y10655 G.gallus NK	801	45	2.6	9487	5	E08264	E08264 DNA fragmen
729	45	2.6					802	45	2.6				
730	45	2.6					803	45	2.6				
731	45	2.6					804	45	2.6				

C 805	45	2.6	90	5	E12579	E12579	Probe. 6/19	878	45	2.6	899	7	BYNCALMOD	L14071 Bryonia dio
C 805	45	2.6	90	5	E12580	E12580	Probe. 6/19	879	45	2.6	742	7	CACORAT6	X71000 C.arellana
C 807	45	2.6	1849	5	E13052	E13052	CDNA. encodl	880	45	2.6	1261	7	CAR012751	AJ012751 Cicer ari
C 808	45	2.6	90	5	E13639	E13639	Probe. 6/19	C 881	45	2.6	2795	7	CBKPMETGN	Z49150 C.blumei K1
809	45	2.6	896	5	I00734	I00734	Sequence 3	882	45	2.6	630	7	CHH111	X77253 C.herbarum
810	45	2.6	2255	5	I04527	I04527	Sequence 1	883	45	2.6	2594	7	CMAC09	X82840 C.melo mRNA
811	45	2.6	1964	5	I08158	I08158	Sequence 6	884	45	2.6	1395	7	CMG1B	X73314 C.maxima MR
812	45	2.6	1964	5	I08168	I08168	Sequence 12	885	45	2.6	1772	7	CREATPS	J04219 C.reinhardt
813	45	2.6	875	5	I13140	I13140	Sequence 1	886	45	2.6	1827	7	CRO7332	AJ007332 Cathartar
814	45	2.6	7859	5	I13185	I13185	Sequence 4	887	45	2.6	774	7	CRRPS18	X83693 C.reinhardt
815	45	2.6	2158	5	I17544	I17544	Sequence 1	888	45	2.6	740	7	CRUB13	X60826 C.reinhardt
816	45	2.6	2010	5	I17767	I17767	Sequence 4	889	45	2.6	2145	7	CSNADP	X58832 C.sorokfina
817	45	2.6	953	5	I19705	I19705	Sequence 1	890	45	2.6	1375	7	EGMDH7	X78800 E.gunnii MR
818	45	2.6	1493	5	I19901	I19901	Sequence 24	891	45	2.6	3271	7	FSU51671	US1671 Fusarium SO
819	45	2.6	2114	5	I22020	I22020	Sequence 6	892	45	2.6	969	7	GHY9725	AJ009725 Gerbera h
820	45	2.6	972	5	I22272	I22272	Sequence 1	893	45	2.6	1754	7	HVADH1	XJ07774 Barley mRNA
821	45	2.6	879	5	I25368	I25368	Sequence 1	894	45	2.6	3292	7	IPBSPP	M64362 Sweet potat
822	45	2.6	3200	5	I26205	I26205	Sequence 1	895	45	2.6	1929	7	LEGALX	X99851 L.erecta MR
823	45	2.6	630	5	I26207	I26207	Sequence 1	896	45	2.6	957	7	LERMALX	X79338 L.esculentu
824	45	2.6	3200	5	I26667	I26667	Sequence 1	897	45	2.6	1032	7	LETHM18GE	X98308 L.esculentu
825	45	2.6	857	5	I29004	I29004	Sequence 1	898	45	2.6	923	7	LEUBC1	X82938 L.esculentu
826	45	2.6	2384	5	I30334	I30334	Sequence 1	899	45	2.6	2142	7	MCMOD1	X64434 M.cystallil
827	45	2.6	2381	5	I30339	I30339	Sequence 6	900	45	2.6	737	7	MTY15368	Y15368 Medicago tr
828	45	2.6	1534	5	I33391	I33391	Sequence 9	901	45	2.6	5617	7	MZEFEGLU	M59190 Zea mays fe
C 829	45	2.6	50	5	I36502	I36502	Sequence 1	C 902	45	2.6	859	7	NPSAPHPRO	X61664 N.sylvestri
C 830	45	2.6	1181	5	I38457	I38457	Sequence 27	C 903	45	2.6	1657	7	NTACCS	X98492 Nicotiana t
C 831	45	2.6	1493	5	I40996	I40996	Sequence 24	C 904	45	2.6	1560	7	NTGPRIBET	X98161 N.tabacum m
C 832	45	2.6	1023	5	I41145	I41145	Sequence 16	C 905	45	2.6	1724	7	NTNKL1RN	X77763 N.tabacum N
C 833	45	2.6	1046	5	I44073	I44073	Sequence 4	C 906	45	2.6	921	7	NTSPAPHP1	X61665 N.tabacum P
C 834	45	2.6	98	5	I49625	I49625	Sequence 42	C 907	45	2.6	1219	7	NTR1BRDR2	X92443 N.tabacum m
C 835	45	2.6	2431	5	I50104	I50104	Sequence 25	C 908	45	2.6	1138	7	NTT0B	X76188 N.tabacum t
C 836	45	2.6	6671	5	I50973	I50973	Sequence 1	C 909	45	2.6	1516	7	OSABATUB1	X78142 O.sativa (A
C 837	45	2.6	1897	5	I52013	I52013	Sequence 1	C 910	45	2.6	1160	7	OSENDO	X56063 O.sativa MR
C 838	45	2.6	1181	5	I56932	I56932	Sequence 27	C 911	45	2.6	1333	7	OSHOX1	X96681 O.sativa MR
C 839	45	2.6	1181	5	I59798	I59798	Sequence 27	C 912	45	2.6	1685	7	OSTA274	X91608 O.sativa MR
C 840	45	2.6	2223	5	I61429	I61429	Sequence 1	C 913	45	2.6	858	7	OSU74296	U74296 Oryza sativ
C 841	45	2.6	536	5	I63075	I63075	Sequence 1	C 914	45	2.6	1148	7	OSY15009	Y15009 Oryza sativ
C 842	45	2.6	2634	5	I66342	I66342	Sequence 1	C 915	45	2.6	800	7	PCPR2	X58698 P.crispum P
C 843	45	2.6	6671	5	I69486	I69486	Sequence 1	C 916	45	2.6	2474	7	PEA70HSP	L03399 Pisum sativ
C 844	45	2.6	1181	5	I75125	I75125	Sequence 27	C 917	45	2.6	1525	7	PEAPAT1	D86180 Pisum sativ
C 845	45	2.6	742	5	I77092	I77092	Sequence 12	C 918	45	2.6	1656	7	PSHEMCA	X73418 P.sativum h
C 846	45	2.6	208	5	I80064	I80064	Sequence 37	C 919	45	2.6	1856	7	PSHEMCA	X73418 P.sativum h
C 847	45	2.6	1425	5	I81228	I81228	Sequence 1	C 920	45	2.6	997	7	PSHEMCA	X73418 P.sativum h
C 848	45	2.6	1425	5	I82208	I82208	Sequence 1	C 921	45	2.6	1596	7	PSHEMCA	X73418 P.sativum h
C 849	45	2.6	857	5	I82533	I82533	Sequence 1	C 922	45	2.6	2036	7	PSY17796	X57429 Pisum sativ
C 850	45	2.6	2158	5	I82810	I82810	Sequence 1	C 923	45	2.6	1776	7	RCEHOLASE	Z28386 R.communis
C 851	45	2.6	117	5	I89931	I89931	Sequence 3	C 924	45	2.6	754	7	RICOREB	D10600 Oryza sativ
C 852	45	2.6	69	5	I89934	I89934	Sequence 7	C 925	45	2.6	764	7	SBANSSPRT	Y09874 S.bicolor m
C 853	45	2.6	69	5	I89944	I89944	Sequence 22	C 926	45	2.6	1768	7	SCSACPD	X78935 S.comerson
C 854	45	2.6	2233	5	I92592	I92592	Sequence 1	C 927	45	2.6	541	7	SHY16247	Y16247 Striga herm
C 855	45	2.6	2384	5	I96058	I96058	Sequence 1	C 928	45	2.6	1132	7	SLU39059	U39059 Lycopersico
C 856	45	2.6	2384	5	I96053	I96053	Sequence 9	C 929	45	2.6	2050	7	SOAHRI	X57073 S.oleracea
C 857	45	2.6	1325	5	I96214	I96214	Sequence 51	C 930	45	2.6	999	7	SOTHRGPB	L22030 Soybean hyd
C 858	45	2.6	723	5	AB015760	AB015760	Nicotiana	C 931	45	2.6	2070	7	SPLAI	X79705 S.pombe pla
C 859	45	2.6	1364	7	AB025573	AB025573	Nicotiana	C 932	45	2.6	599	7	TAEPM1	X68289 T.aestivum
C 860	45	2.6	1312	7	AB025715	AB025715	Nicotiana	C 933	45	2.6	1174	7	TAVDAC2	X82149 T.aestivum
C 861	45	2.6	1978	7	AOASPSYNM	AOASPSYNM	A.officinal	C 934	45	2.6	565	7	TDDH25	X78430 T.durum Des
C 862	45	2.6	2190	7	ATAXAA	ATAXAA	A.thaliana	C 935	45	2.6	2564	7	TREX11	Z69257 T.reesei MR
C 863	45	2.6	1636	7	ATASKDZET	ATASKDZET	A.thaliana	C 936	45	2.6	2247	7	VFA011303	AJ011303 Vicia fab
C 864	45	2.6	1372	7	ATCOQ3	ATCOQ3	Arabidopsis	C 937	45	2.6	1146	7	VORNACH11	X88600 V.unguicula
C 865	45	2.6	1364	7	ATCSMRNA	ATCSMRNA	A.thaliana	C 938	45	2.6	972	7	VURNACH14	X88803 V.unguicula
C 866	45	2.6	1549	7	ATH7450	ATH7450	Arabidops	C 939	45	2.6	7805	7	YSCC02INDU	L19873 Yeast pYEU1
C 867	45	2.6	1414	7	ATH7587	ATH7587	Arabidops	C 940	45	2.6	918	7	ZMANB1B	X66077 Z.mays MNB1
C 868	45	2.6	1462	7	ATHRPCA	ATHRPCA	Arabidopsis	C 941	45	2.6	1048	7	ZMPRO1MR	X63379 Z.mays ZMR1
C 869	45	2.6	2923	7	ATHSP91	ATHSP91	A.thaliana	C 942	45	2.6	1065	7	ZMYPT785	X63378 Z.mays yptm
C 870	45	2.6	753	7	ATP29A	ATP29A	A.thaliana	C 943	45	2.6	1725	8	AF007785	AF007785 Zea mays
C 871	45	2.6	940	7	ATSTZ	ATSTZ	A.thaliana	C 944	45	2.6	946	8	AF010228	AF010228 Lycopersi
C 872	45	2.6	1080	7	ATTIIPRO	ATTIIPRO	A.thaliana	C 945	45	2.6	1848	8	AF012862	AF012862 Petrosell
C 873	45	2.6	2357	7	BNAP609	BNAP609	Brassica	C 946	45	2.6	808	8	AF022741	AF022741 Oryza sat
C 874	45	2.6	2085	7	BOMASPCN	BOMASPCN	X84448 B.oleracea	C 947	45	2.6	1784	8	AF029856	AF029856 Sorghum b
C 875	45	2.6	1114	7	BRUWIPBICP	BRUWIPBICP	B.oleracea	C 948	45	2.6	1319	8	AF031609	AF031609 Oryza sat
C 876	45	2.6	578	7	BRURIRIPL	BRURIRIPL	L21898 Brassica ra	C 949	45	2.6	1380	8	AF038585	AF038585 Zea mays
C 877	45	2.6	567	7	BRUBIRBPA	BRUBIRBPA	Z24738 B.rapa ubiq	C 950	45	2.6	806	8	AF038949	AF038949 Pinus mon

951	45	2.6	2233	8	AF049853	AF049853	Neurospor
952	45	2.6	591	8	AF049919	AF049919	Petunia x
953	45	2.6	423	8	AF049929	AF049929	Petunia x
954	45	2.6	1474	8	AF056283	AF056283	Nicotiana
955	45	2.6	1646	8	AF058708	AF058708	Zea mays
956	45	2.6	2503	8	AF061107	AF061107	Zea mays
957	45	2.6	2780	8	AF064561	AF064561	Hordeum v
958	45	2.6	1049	8	AF071889	AF071889	Prunus ar
959	45	2.6	981	8	AF082032	AF082032	Hemerocal
960	45	2.6	734	8	AF088913	AF088913	Petunia x
961	45	2.6	845	8	AF089846	AF089846	Funaria h
962	45	2.6	2363	8	AF090835	AF090835	Mesembrya
963	45	2.6	1151	8	AF095927	AF095927	Secale ce
964	45	2.6	3916	8	AF097728	AF097728	Aspergill
965	45	2.6	1446	8	AF106660	AF106660	Lycopersi
966	45	2.6	2534	8	AF117125	AF117125	Arabidops
967	45	2.6	1008	8	AF117707	AF117707	Lycopersi
968	45	2.6	3336	8	AF118843	AF118843	Lycopersi
969	45	2.6	1057	8	AF145730	AF145730	Oryza sat
970	45	2.6	1559	8	AF082830	AF082830	Oryza sat
971	45	2.6	1364	8	ANU51131	ANU51131	Aspergillus
972	45	2.6	108484	8	ATAC004238	ATAC004238	Arabidops
973	45	2.6	86686	8	ATAC004786	ATAC004786	Arabidops
974	45	2.6	118196	8	ATAC005824	ATAC005824	Arabidops
975	45	2.6	907	8	ATU95035	ATU95035	Arabidops
976	45	2.6	1407	8	BRU223455	BRU223455	Brassica
977	45	2.6	1710	8	BRU223498	BRU223498	Brassica
978	45	2.6	1136	8	CMU61386	CMU61386	Cucurbita m
979	45	2.6	1047	8	CRU89649	CRU89649	Chlamydomon
980	45	2.6	737	8	DAREP21P	DAREP21P	Daucus caro
981	45	2.6	3981	8	ENU01146	ENU01146	Emeticella
982	45	2.6	3284	8	LEU86662	LEU86662	Lycopersico
983	45	2.6	1305	8	PAU82330	PAU82330	Prunus arme
984	45	2.6	636	8	PAU95179	PAU95179	Prunus arme
985	45	2.6	691	8	PAU97494	PAU97494	Prunus arme
986	45	2.6	890	8	S63825	S63825	NADPH-proto
987	45	2.6	1364	8	S83229	S83229	Solenum com
988	45	2.6	815	8	SCU92087	SCU92087	Solenum com
989	45	2.6	1044	9	AB019565	AB019565	Homo sapi
990	45	2.6	100000	9	AP000020	AP000020	Homo sapi
991	45	2.6	100000	9	AP000161	AP000161	Homo sapi
992	45	2.6	84126	9	HS125N5	HS125N5	Human DNA
993	45	2.6	91835	9	HS26H23	HS26H23	Human DNA
994	45	2.6	149486	9	HS427A4	HS427A4	Human DNA
995	45	2.6	151213	9	HS441J1	HS441J1	Human DNA
996	45	2.6	118226	9	HS46618	HS46618	Human DNA
997	45	2.6	199016	9	HS48612	HS48612	Human DNA
998	45	2.6	109859	9	HS747E2	HS747E2	Human sapi
999	45	2.6	139904	9	HS97D16	HS97D16	Homo sapi
1000	45	2.6	4281	9	HSA012288	HSA012288	Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AR015970 1737 bp DNA
DEFINITION Sequence 3 from patent US 5776457.
ACCESSION AR015970
NID 93972247
VERSION AR015970.1 GI:3972247
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1737)
AUTHORS Lee, J. and Wood, W.I.
TITLE Antibodies to human PFA4 receptor and compositions thereof
JOURNAL Patent: US 5776457-A 3 07-JUL-1998;
FEATURES Location/Qualifiers
1..1737
/organism="unknown"

PAT 04-DEC-1998

BASE COUNT	454 a	411 c	373 g	499 t
ORIGIN				
Query Match	100.0%;	Score 1737;	DB 5;	Length 1737;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1737; Conservative	0;	Mismatches	0;	Indels 0; Gaps
QY	1	GAATTCAGTGTGCTGGGCGGCGGCGCAAGTGCACCGGAGGGCCCTGAGTCTCAGTA	60	
DB	1	GAATTCAGTGTGCTGGGCGGCGGCGCAAGTGCACCGGAGGGCCCTGAGTCTCAGTA	60	
QY	61	GCCACCGCATCTGAGAACACGCGTTACCATGAGGGGATCATATATACATTCAGAT	120	
DB	61	GCCACCGCATCTGAGAACACGCGTTACCATGAGGGGATCATATATACATTCAGAT	120	
QY	121	AACATACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGAACCTGTTCCGT	180	
DB	121	AACATACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAGAACCTGTTCCGT	180	
QY	181	GAGGAAATGCTAATTTCAATAAATCTTCGCGCCACATCTACTCCATCTTTCTTA	240	
DB	181	GAGGAAATGCTAATTTCAATAAATCTTCGCGCCACATCTACTCCATCTTTCTTA	240	
QY	241	ACTGCATTTGGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG	300	
DB	241	ACTGCATTTGGGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATG	300	
QY	301	AGCATGACGAGCAAGTACAGGCTGCACCTGATGAGCCGACCTCTTTGTCATCAG	360	
DB	301	AGCATGACGAGCAAGTACAGGCTGCACCTGATGAGCCGACCTCTTTGTCATCAG	360	
QY	361	CTTCCCTTCTGGGCGAGTTGATGCCGTGGCAACTGGTATTTGGGAATCTTCATCAG	420	
DB	361	CTTCCCTTCTGGGCGAGTTGATGCCGTGGCAACTGGTATTTGGGAATCTTCATCAG	420	
QY	421	GCATTCATGTCATCTACACAGTCAACCTCTACACAGTCTGATGATGATGATGATG	480	
DB	421	GCATTCATGTCATCTACACAGTCAACCTCTACACAGTCTGATGATGATGATGATG	480	
QY	481	AGCTGAGCGCTACTGCGCATCTGTCACAGCCACCAAGTACAGGCGCAAGAGCTG	540	
DB	481	AGCTGAGCGCTACTGCGCATCTGTCACAGCCACCAAGTACAGGCGCAAGAGCTG	540	
QY	541	TTGGCTGAAAGAGTGTATGTTGGCTGTGATGATGATGATGATGATGATGATGATG	600	
DB	541	TTGGCTGAAAGAGTGTATGTTGGCTGTGATGATGATGATGATGATGATGATGATG	600	
QY	601	GACTTCATCTTTGCCACGTCAGTGAAGGACATGACAGATATATCTGTGACCGCTTAC	660	
DB	601	GACTTCATCTTTGCCACGTCAGTGAAGGACATGACAGATATATCTGTGACCGCTTAC	660	
QY	661	CCCATGACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	720	
DB	661	CCCATGACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	720	
QY	721	CCTGATATGTCATCT	780	
DB	721	CCTGATATGTCATCT	780	
QY	781	GCCACACGAAAGCGAAGGCGCCCTCAAGACACAGTCACTCTGCTGCTTCTTGCC	840	
DB	781	GCCACACGAAAGCGAAGGCGCCCTCAAGACACAGTCACTCTGCTGCTTCTTGCC	840	
QY	841	TGTTGGCTGCTTACTATGATGGGATGATGATGATGATGATGATGATGATGATGATG	900	
DB	841	TGTTGGCTGCTTACTATGATGGGATGATGATGATGATGATGATGATGATGATGATG	900	
QY	901	AAGCAAGGCTGTGAGTTGAGAACACTGTGCACAGTGAATTTCCATCAGGAGCCCTA	960	
DB	901	AAGCAAGGCTGTGAGTTGAGAACACTGTGCACAGTGAATTTCCATCAGGAGCCCTA	960	
QY	961	GCTTCTTCCACTGTTGTCTGAACCCCATCTCTATGCTTCTTGAGCCAAATTTAA	1020	

Db	961	GCCTTCCTCCACAGTGTGCTGTAACCCCACTCCTCATAGCTTTCCTTGGACCACAAATTTAA	1020
Qy	1021	ACCTTGTGCCACGACGCACATCTACTCTGTGTAGACGAGGTGCCAGCTCAAGATTCCTCTCC	1080
Db	1021	ACCTGTGCCACGACGCACACTCCCTCTGTGAGCAGAGGGTCCAGCCTCAAGATTCCTCTCC	1080
Qy	1081	AAAGGAAGCGCAGGTGGACATCTACTCTGTTCACAGTCGAGTCGAGTCTTCAGATTTTAC	1140
Db	1081	AAAGGAAGCGCAGGTGGACATCTACTCTGTTCACAGTCGAGTCGAGTCTTCAGATTTTAC	1140
Qy	1141	TCCACCTTAACACAGATGTAAAAAGACTTTTATATAGATTAATAACTTTTATTAAGTT	1200
Db	1141	TCCACCTTAACACAGATGTAAAAAGACTTTTATATAGATTAATAATACTTTTATTAAGTT	1200
Qy	1201	ACACATTTTGTAGATATATAAAGACTGACCAATATTGTACAGTTTTTATGCTTGTGGAT	1260
Db	1201	ACACATTTTGTAGATATATAAAGACTGACCAATATTGTACAGTTTTTATGCTTGTGGAT	1260
Qy	1261	TTTTCTCTGTGTTCTCTTGTAGTTTTTGTGAGTTTATGACTTATTTATATAATTTTT	1320
Db	1261	TTTTCTCTGTGTTCTCTTGTAGTTTTTGTGAGTTTATGACTTATTTATATAATTTTT	1320
Qy	1321	TTTGTTCATATGTAGTGTGTCTAGCGAGACCTGTGGCCAAAGTCTTAGTGTCTGAT	1380
Db	1321	TTTGTTCATATGTAGTGTGTCTAGCGAGACCTGTGGCCAAAGTCTTAGTGTCTGAT	1380
Qy	1381	GTCTCTGTGTAGACTGTAGAAAAAGGAACTGAACTTCCAGAGGTGTAGTGAATCAG	1440
Db	1381	GTCTCTGTGTAGACTGTAGAAAAAGGAACTGAACTTCCAGAGGTGTAGTGAATCAG	1440
Qy	1441	TAAAGCTGAATATATATCCCGAGCTGTTTATGCATGATATATCTTCACATCCCGTGGAA	1500
Db	1441	TAAAGCTGAATATATATCCCGAGCTGTTTATGCATGATATATCTTCACATCCCGTGGAA	1500
Qy	1501	GTTTTTCTGTCTTCTTAAGACGTGATTTTGTCTGTAGAAGATGGACACTTAATAACCAAGGCC	1560
Db	1501	GTTTTTCTGTCTTCTTAAGACGTGATTTTGTCTGTAGAAGATGGACACTTAATAACCAAGGCC	1560
Qy	1561	AAAGGTGATGAAGATCTGCTGTTTTCAGTTTCAAGAGTGGGTGATTTACGACCTTAC	1620
Db	1561	AAAGGTGATGAAGATCTGCTGTTTTCAGTTTCAAGAGTGGGTGATTTACGACCTTAC	1620
Qy	1621	AGTGTACAGTCTTGTATTAGTGTATTAATAAAGTACATGTAAACTTAATAAAAAAAAAA	1680
Db	1621	AGTGTACAGTCTTGTATTAGTGTATTAAAGTACATGTAAACTTAATAAAAAAAAAA	1680
Qy	1681	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACGACACTGGAATTC	1737
Db	1681	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACGACACTGGAATTC	1737
RESULT	2		
LOCUS	113753	1737 bp	DNA
DEFINITION	Sequence 4	from patent US 5440021.	PAT
ACCESSION	113753		
NTD	g996819		
VERSION	113753.1	GI:996819	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1737)		
AUTHORS	Chuntharapai,A., Hebert,C., Kim,K.J. and Lee,J.		
TITLE	Antibodies to human IL-8 type B receptor		
JOURNAL	Patent: US 5440021-A 4 08-AUG-1995;		
FEATURES	Location/Qualifiers		
source	1..1737	/organism="unknown"	
BASE COUNT	454 a	411 c	373 g
ORIGIN			499 t

[illegible]

```

OY 1081 AAGGAAGGAGGAGGATTCATCTGTTCCAGTGTAGTCTCAGTTTTCAC 1140
    |||||||
Db 1081 AAGGAAGGAGGAGGATTCATCTGTTCCAGTGTAGTCTCAGTTTTCAC 1140
OY 1141 TCCAGCTAACACAGATGTAAAGACTTTTATACGATTAATACCTTTTAAAGTT 1200
    |||||||
Db 1141 TCCAGCTAACACAGATGTAAAGACTTTTATACGATTAATACCTTTTAAAGTT 1200
OY 1201 AACCATTTTTCACATATAAAGACTGACCAATATGTACAGTTTATGCTTTGGAT 1260
    |||||||
Db 1201 AACCATTTTTCACATATAAAGACTGACCAATATGTACAGTTTATGCTTTGGAT 1260
OY 1261 TTTTGTCTGTCTTTCTTAAGTTTGTGAAGTTTAATGACTTATTAATAATTTT 1320
    |||||||
Db 1261 TTTTGTCTGTCTTTCTTAAGTTTGTGAAGTTTAATGACTTATTAATAATTTT 1320
OY 1321 TTTTGTCTGTCTTTCTTAAGTTTGTGAAGTTTAATGACTTATTAATAATTTT 1380
    |||||||
Db 1321 TTTTGTCTGTCTTTCTTAAGTTTGTGAAGTTTAATGACTTATTAATAATTTT 1380
OY 1381 GTCTCGGTAGACGCTGTAGAAAAGGAATTCACAGACCGTGTAGTGAATCAG 1440
    |||||||
Db 1381 GTCTCGGTAGACGCTGTAGAAAAGGAATTCACAGACCGTGTAGTGAATCAG 1440
OY 1441 TAAAGCTAGAAATGATCCCGAGCTGTTATGATAGATTAATCTCCATTCGCTGGAAC 1500
    |||||||
Db 1441 TAAAGCTAGAAATGATCCCGAGCTGTTATGATAGATTAATCTCCATTCGCTGGAAC 1500
OY 1501 GTTTTCTCTTCTTTAAGACGTGATTTTGTCTGTAAGAAATGGCACTTATAACCAAGCCC 1560
    |||||||
Db 1501 GTTTTCTCTTCTTTAAGACGTGATTTTGTCTGTAAGAAATGGCACTTATAACCAAGCCC 1560
OY 1561 AAGTGGTATAGAAAGCTGTTTTCAGTTTCAGAGAGGGTGTATTCAGACCTAC 1620
    |||||||
Db 1561 AAGTGGTATAGAAAGCTGTTTTCAGTTTCAGAGAGGGTGTATTCAGACCTAC 1620
OY 1621 AGTGTACAGTCTGTATTAAGTGTATTAAGTAAAGTACATGTAAACTTAAAAA 1680
    |||||||
Db 1621 AGTGTACAGTCTGTATTAAGTGTATTAAGTAAAGTACATGTAAACTTAAAAA 1680
OY 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACACACTGGAATTC 1737
    |||||||
Db 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACACACTGGAATTC 1737

```

```

RESULT 4
LOCUS HSNPRLA 1645 bp mRNA PRI 15-FEB-1994
DEFINITION H.sapiens mRNA for neuroleptide Y-like receptor.
ACCESSION X71635
NID 9297099
VERSION X71635.1 GI:297099
KEYWORDS Chemotaxis; G-protein coupled receptor; neuroleptide Y.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1645)
AUTHORS Moser, B.
TITLE Direct Submission
JOURNAL Submitted (26-APR-1993) B. Moser, University of Bern, PO Box 99,
Bern, SWITZERLAND
2 (bases 1 to 1645)
AUTHORS Moser, B.
TITLE Cloning of a human seven-transmembrane domain receptor, LESTR, that
is highly expressed in leukocytes
JOURNAL J. Biol. Chem. 269 (1), 232-237 (1994)
MEDLINE 94103215
FEATURES
Location/Qualifiers
1..1645
/organism="Homo sapiens"

```

```

BASE COUNT 402 a 395 c 356 g 492 t
ORIGIN
Query Match 94.5%; Score 1642; DB 10; Length 1645;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


Db	664	CCAGTTTCAGACATCATGTTGGCCCTTATCTCGCCGATTTGTCATCTCTGCTCCGTA	723
Qy	747	TTGCATTATCATCTCCAAAGCTGTGCACCTCCAAAGGGCCACAGAGCCGAAAGCCCTCAA	806
Db	724	TTGCATTATCATCTCCAAAGCTGTGCACCTCCAAAGGGCCACAGAGCCGAAAGCCCTCAA	783
Qy	807	GACCAAGCTACCTCCATCCGCGGTTCTTGGCTGTGGTCCCTTACATCTGGGAT	866
Db	784	GACCACATCATCTCCATCCGCGGTTCTTGGCTGTGGTCCCTTACATCTGGGAT	843
Qy	867	CAGCATGACCTCTTCATCCCTCGAAGATCATCAAGCAAGGGTGTGAGTTTGAGAACAC	926
Db	844	CAGCATGACCTCTTCATCCCTCGAAGATCATCAAGGGTGTGAGTTTGAGAACAC	903
Qy	927	TGTGCACAAGTGGATTTTCATACCCGAGGCCCTAGCTTTCTTCACCTGTTGTGTGACCC	986
Db	904	TGTGCACAAGTGGATTTTCATACCCGAGGCCCTAGCTTTCTTCACCTGTTGTGTGACCC	963
Qy	987	CATCCTCATGCTTTCTCTGGAGCAATTTAAACCTCTGCCACAGCACTCACTCC	1044
Db	964	CATCCTCATGCTTTCTCTGGAGCAATTTAAACCTCTGCCACAGCACTCACTCC	1022
Qy	1047	TGTGAGCAGAGGCTCCAGCCCAAGATCTCTCCAAAGAAACGAGGTGAGCATTCATC	1100
Db	1024	TGTGAGCAGAGGCTCCAGCCCAAGATCTCTCCAAAGAAACGAGGTGAGCATTCATC	1088
Qy	1107	TGTTCCACTGAGTGTGAGTCTTCAAGTTTCACTCCAGCTAAACAGATGTAAAGACT	1166
Db	1084	TGTTCCACTGAGTGTGAGTCTTCAAGTTTCACTCCAGCTAAACAGATGTAAAGACT	1144
Qy	1187	TTTTTTTAAACGATTAATTAATCTTTTAAAGTAAACATTTTTCAGATTAATAAGACTG	1222
Db	1144	TTTTTTTAAACGATTAATTAATCTTTTAAAGTAAACATTTTTCAGATTAATAAGACTG	1200
Qy	1227	ACCAATATGTACAGTTTATATGCTTGTGAGATTTTGTCTGTGCTTTCTTAAAGTTT	1286
Db	1204	ACCAATATGTACAGTTTATATGCTTGTGAGATTTTGTCTGTGCTTTCTTAAAGTTT	1264
Qy	1287	GTGAAGTTTAATGACTTATTTATATAAATTTTTTGTTCATATTAATGATGTGTCTAG	1344
Db	1264	GTGAAGTTTAATGACTTATTTATATAAATTTTTTGTTCATATTAATGATGTGTCTAG	1322
Qy	1347	GCAGAGCTGTGGCCAGTTCTTATGTCGTGATGTCCTGTGTAGACTGTAGAAAG	1400
Db	1324	GCAGAGCTGTGGCCAGTTCTTATGTCGTGATGTCCTGTGTAGACTGTAGAAAG	1388
Qy	1407	GACGTAACATTCAGAGCGGTGTGAATCAGGTAAAGCTAAGTAAATGCCACGCT	1466
Db	1384	GACGTAACATTCAGAGCGGTGTGAATCAGGTAAAGCTAAGTAAATGCCACGCT	1444
Qy	1467	TTATGATAGATAATCTCTCCATCCCGGTGAAGAGTTTTCCTGTTCTTAAGACGTAT	1522
Db	1444	TTATGATAGATAATCTCTCCATCCCGGTGAAGAGTTTTCCTGTTCTTAAGACGTAT	1500
Qy	1527	TTTGTGTGAAGATGCGACTTATTAACAAAGCCCAAGTGTATAAATGCTGGTTTT	1586
Db	1504	TTTGTGTGAAGATGCGACTTATTAACAAAGCCCAAGTGTATAAATGCTGGTTTT	1564
Qy	1587	CAGTTTACAGAGTGGGTGATTTACACACCTACAGTACAGTCTGATTAAGTGT	1646
Db	1564	CAGTTTACAGAGTGGGTGATTTACACACCTACAGTACAGTCTGATTAAGTGT	1622
Qy	1647	AATAAAGTACATGTTAAACTT 1668	
Db	1624	AATAAAGTACATGTTAAACTT 1645	
RESULT	5		
HUMHM89			
LOCUS	HUMHM89	1664 bp	mrna
DEFINITION	Human mRNA for HM89.		
ACCESSION	D10924		
MD	g219868		
			03-FEB-1999

KEYWORDS	D10924.1 GI:219868
FEATURES	GTP-binding protein; plasma membrane protein; protein coupled.
SOURCE	Homo sapiens monocyte, cDNA to mRNA.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
AUTHORS	Primates; Catarrhini; Homidae; Homo.
TITLE	1 (bases 1 to 1664)
JOURNAL	Nomura, H., Nielsen, B.W. and Matsushima, K.
MEDLINE	Molecular cloning of cDNAs encoding a LD78 receptor and putative leukocyte chemotactic peptide receptors
COMMENT	Int. Immunol. 5 (10), 1239-1249 (1993) 94092629 Submitted (13-Apr-1992) to DDBJ by: Hideki Nomura Dept. Of Pharmacol. Cancer Res. Inst., Kanazawa Univ. 13-1 Takaramachi Kanazawa, Ishikawa 920 Japan Phone: 0762-62-8151 x5875 Fax: 0762-60-7704.
FEATURES	Location/Qualifiers
source	1..1664 /organism="Homo sapiens" /db_xref="taxon:9606" /cell_line="human monocyte" /cell_type="monocyte" 77..1135 /gene="HM89" 77..1135 /gene="HM89" /codon_start=1 /product="HM89" /protein_id="BAA01722.1" /db_xref="PID:d1002197" /db_xref="PID:g219869" /db_xref="GI:219869" /translation="MESISYTSNTNTEEMSGDYDMKPECPREENANFKFLPII YSIFLTVIGNGVLIVIMVGYOKRLRMTQRYRLHSVALDLEVTILPEWADVAVNM WFGFLCKAHHVITVNLVSDDLILAFISDRYLAIHVHANSGRPRLAEKVYGV WIPLLLTIDPFIANYSEADRYICDFEYPNDLMWPYFQFOHMYGLILPGVIIISC YCIIKSLSKGKORAKLTYYIIILAFACMLPYIISIDSFTLEIKOGCEFF ENLYHKMISTEALAFPHCLNPILYVFLGAKRTSNQHALTYSRSSSLIKLSGRK GGHSSTSESSSFHS"
BASE COUNT	414 a 398 c 359 g 493 t
ORIGIN	
Query Match	92.28; Score 1601; DB 9; Length 1664; Best Local Similarity 99.99; Pred. No. 0; Matches 1651; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy	27 GCAAAGTGAGCCCGAGGAGCTGAGTGTCCAGTAGGCCACCGCATCTGAGAACCAGCGGT 86
Db	13 GCAAGTAGACCCCGAGGAGCTGAGTGTCCAGTAGGCCACCGCATCTGAGAACCAGCGGT 72
Oy	87 TACCATGGAGGGGATCAGTATATACACTTCAGATACTACACGAGGAATGGGCTCAGG 146
Db	73 TACCATGGAGGGGATCAGTATATACACTTCAGATACTACACGAGGAATGGGCTCAGG 132
Oy	147 GGACATATATCTTCATGAAGGAACCTGTTCCTCGTGAGAAAAATGCTAATTTCATAAAT 206
Db	133 GGACATATATCTTCATGAAGGAACCTGTTCCTCGTGAGAAAAATGCTAATTTCATAAAT 192
Oy	207 CTTCCTGGCCACATCTATCTGCATCATCTTTCTTAACGTGGCATTTGGGCAATGGATTGGT 266
Db	193 CTTCCTGGCCACATCTATCTGCATCATCTTTCTTAACGTGGCATTTGTGGCAATGGATTGGT 252
Oy	267 CATCTGTGTCATGGGTTACCAAGAAGAACTGAGAAGCATGACGGACAAGTACAGGCTGCA 326
Db	253 CATCTGTGTCATGGGTTACCAAGAAGAACTGAGAAGCATGACGGACAAGTACAGGCTGCA 312
Oy	327 CCTGTCACTGGCCGACCTCCTCTTTGTCATCACGCTTCCTTGTGGGCAATTGATGCCGT 386

STS /map="2"
 1220. .1426
 primer_bind 1220. .1244
 complement(1405. .1426)
 BASE COUNT 414 a 398 c 359 g 493 t
 ORIGIN

Query Match 92.2% Score 1601; DB 14; Length 1664;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1651; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 27 GCAAGTACGCGGAGGCTGAGTGTCCAGTAGCCAGCGCATCTGGAGAACCGGCT 86
 Db 13 GCAAGTACGCGGAGGCTGAGTGTCCAGTAGCCAGCGCATCTGGAGAACCGGCT 72
 OY 87 TACCATGAGGAGTCACTATATACACTCAGATTAACCCAGAGAAATGGCTAGG 146
 Db 73 TACCATGAGGAGTCACTATATACACTCAGATTAACCCAGAGAAATGGCTAGG 132
 OY 147 GACATATGATCCATGAGAGAACCTGTTCCGAGAGAAATGCTAATTTCAATAAAT 206
 Db 133 GACATATGATCCATGAGAGAACCTGTTCCGAGAGAAATGCTAATTTCAATAAAT 192
 OY 207 CTTCTGCCACCATCTACTTCATCATCTTCTTAAGTGGCATTGGGCAATGGATTGT 266
 Db 193 CTTCTGCCACCATCTACTTCATCATCTTCTTAAGTGGCATTGGGCAATGGATTGT 252
 OY 267 CATCTGCTGATGGGTACCAAGAGAAAGTGAAGAGATGACGACACAGTACAGCTGCA 326
 Db 253 CATCTGCTGATGGGTACCAAGAGAAAGTGAAGAGATGACGACACAGTACAGCTGCA 312
 OY 327 CTTCTGAGTGGCGACCTCTCTTTCATCAGCTCTCCCTTCGGGAGCTGATGGCGT 386
 Db 313 CTTCTGAGTGGCGACCTCTCTTTCATCAGCTCTCCCTTCGGGAGCTGATGGCGT 372
 OY 387 GGGAACTGCTACTTTGGGAATCTCTATGCAAGGAGTCCATGTCATCTACACAGTCAA 446
 Db 373 GGGAACTGCTACTTTGGGAATCTCTATGCAAGGAGTCCATGTCATCTACACAGTCAA 432
 OY 447 CCTCTACAGAGTGTCTCTATCCTGGGCTTCATCAGTCTGGACCGCTACTGGGCATCGT 506
 Db 433 CCTCTACAGAGTGTCTCTATCCTGGGCTTCATCAGTCTGGACCGCTACTGGGCATCGT 492
 OY 507 CCACGCCACCAAGTACAGAGGCCCAAGAGAGCTGTTGGTGAAGAGTGTCTATGTTGG 566
 Db 493 CCACGCCACCAAGTACAGAGGCCCAAGAGAGCTGTTGGTGAAGAGTGTCTATGTTGG 552
 OY 567 CGTCTGGATCCCTGCTCTCTGCTGACTATTCGGACTTATCTTTGCCAAGCTCACTGA 626
 Db 553 CGTCTGGATCCCTGCTCTCTGCTGACTATTCGGACTTATCTTTGCCAAGCTCACTGA 612
 OY 627 GCGAGATGAGAGATATATCTGTGACCGCTCTACCCCAATGACTGTGGGTGTGTT 686
 Db 613 GCGAGATGAGAGATATATCTGTGACCGCTCTACCCCAATGACTGTGGGTGTGTT 672
 OY 687 CCAGTTTACAGACATCATGTTGGCTTATCTGCTGCTGATTTGCTATCTGCTGCTGA 746
 Db 673 CCAGTTTACAGACATCATGTTGGCTTATCTGCTGCTGATTTGCTATCTGCTGCTGA 732
 OY 747 TTGCATTATCATCTCAAGCTGTCTACACTCCAAAGGGCCACAGAGGGCAAGGCCCTCAA 806
 Db 733 TTGCATTATCATCTCAAGCTGTCTACACTCCAAAGGGCCACAGAGGGCAAGGCCCTCAA 792
 OY 807 GACACAGTATCTCTATCTGTGCTTCTGTGCTGTGGCTGCTTACTACTATGGGAT 866
 Db 793 GACACAGTATCTCTATCTGTGCTTCTGTGCTGTGGCTGCTTACTACTATGGGAT 852
 OY 867 CACCATGACCTCTCTCTCTCTGGAATCATCAACCAAGGGGTGAGTTTGAGACAC 926
 Db 853 CACCATGACCTCTCTCTCTCTCTGGAATCATCAACCAAGGGGTGAGTTTGAGACAC 912
 OY 927 TGTGCAACAAGTGAATTTTCATCAACGAGGCCCTAGCTTTCTTCCACTGTGTGTGAACCC 986

Db 913 TGTGCAACAAGTGAATTTTCATCAACGAGGCCCTAGCTTTCTTCCACTGTGTGTGAACCC 972
 OY 987 CATCTCTATGCTTTCTCTGGAGCCAAATTTAAACCTCTGCCACAGACTCACTC 1046
 Db 973 CATCTCTATGCTTTCTCTGGAGCCAAATTTAAACCTCTGCCACAGACTCACTC 1032
 OY 1047 TGTGAGAGAGGGTCCAGCCTCAAGATCTCTCCAAAGGAAAGCGAGTGGACATTCATC 1106
 Db 1033 TGTGAGAGAGGGTCCAGCCTCAAGATCTCTCCAAAGGAAAGCGAGTGGACATTCATC 1092
 OY 1107 TGTTCCTGAGTCTGAGTCTTCAAGTTTTCCTCAGCTAACACAGATGTAAAGACT 1166
 Db 1093 TGTTCCTGAGTCTGAGTCTTCAAGTTTTCCTCAGCTAACACAGATGTAAAGACT 1152
 OY 1167 TTTTATATGATTAATAAATTTTAAAGTACACATTTTTCAGATTAATAAGACT 1226
 Db 1153 TTTTATATGATTAATAAATTTTAAAGTACACATTTTTCAGATTAATAAGACT 1212
 OY 1227 ACCAATATGATTAATAAATTTTAAAGTACATTTTTCAGATTAATAAGACT 1286
 Db 1213 ACCAATATGATTAATAAATTTTAAAGTACATTTTTCAGATTAATAAGACT 1272
 OY 1287 GTGAGTTTATGATTAATAAATTTTTCAGATTAATAAGACT 1346
 Db 1273 GTGAGTTTATGATTAATAAATTTTTCAGATTAATAAGACT 1332
 OY 1347 GCAAGACCTGTGGCAGTCTTGTGCTGATGCTGTGCTGTGAGTGTGAAGAG 1406
 Db 1333 GCAAGACCTGTGGCAGTCTTGTGCTGATGCTGTGCTGTGAGTGTGAAGAG 1392
 OY 1407 GAACTGACATTCAGAGGCTGTAGTGAATCAGCTAAGCTAAGATGATCCACACTGT 1466
 Db 1393 GAACTGACATTCAGAGGCTGTAGTGAATCAGCTAAGCTAAGATGATCCACACTGT 1452
 OY 1467 TTAATGATATATCTCTCATTCCCTGGAGACGTTTTCCTGTTTAAAGCGTGAAT 1526
 Db 1453 TTAATGATATATCTCTCATTCCCTGGAGACGTTTTCCTGTTTAAAGCGTGAAT 1512
 OY 1527 TTGCTGAGAGATGGCAGTATTAACCAAGGCCAAGGTTAAGATGCTGGTTTTT 1586
 Db 1513 TTGCTGAGAGATGGCAGTATTAACCAAGGCCAAGGTTAAGATGCTGGTTTTT 1572
 OY 1587 CAGTTTTCAGAGATGGGTTGATTTTCAGACCTACAGTGTACAGTCTTGAATTAAGTTGT 1646
 Db 1573 CAGTTTTCAGAGATGGGTTGATTTTCAGACCTACAGTGTACAGTCTTGAATTAAGTTGT 1632
 OY 1647 AATTAAGTACATGTTAACTTAAAAA 1678
 Db 1633 AATTAAGTACATGTTAACTTAAAAA 1664

RESULT 7
 AF005058
 LOCUS AF005058 5161 bp DNA PRI 01-JAN-1998
 DEFINITION Homo sapiens chemokine receptor (CXCR-4) gene, complete cds.
 ACCESSION AF005058
 NID 92735718
 VERSION AF005058.1 GI:2735718
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 5161)
 AUTHORS Wegner,S.A., Ehrenberg,P.K., Chang,G., Dayhoff,D.E. and Michael,N.L.
 TITLE Genomic organization and characterization of the promoter for the HIV-1 entry co-receptor CXCR-4
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 5161)
 AUTHORS Wegner,S.A., Ehrenberg,P.K., Chang,G., Dayhoff,D.E. and Michael,N.L.

TITLE Direct Submission
JOURNAL Submitted (21-MAY-1997) Division of Retrovirology, Walter Reed Army
Institute of Research, 13 Taft Court, Suite 200, Rockville, MD
20850, USA

FEATURES Location/Qualifiers
source 1. .5161
/organism="Homo sapiens"
/db_xref="taxon:9606"
2. .5161
/gene="CXCR-4"
TATA_signal 919. .923
/gene="CXCR-4"
exon 950. .1052
/gene="CXCR-4"
5'UTR 950. .1037
/number=1
/gene="CXCR-4"
mRNA join(950. .1052,3185. .4760)
/gene="CXCR-4"
/product="chemokine receptor"
join(1038. .1052,3185. .4228)
/gene="CXCR-4"
CDS /note="seven transmembrane G-coupled receptor for the CXCR
chemokine SDF-1; entry co-receptor for HIV-1 strains with
the syncytium-inducing phenotype"
/codon_start=1
/product="chemokine receptor"
/protein_id="AA893982.1"
/db_xref="PID:q2735719"
/db_xref="GI:2735719"
/translation="MEGISLYTSDNYTEKSGSDYSMEKPEFRENNANFNKLFPTI
YSIIFFGIIVNGLVILVMGYOKRLSMDKRYLHLSVADLFVITLPEVAADVAM
YFGNFKCAVHIVLYNLVSVLILAFISLDRLAIYHATNSGRPKLLAEKVVYV
WIPALLITIPDFEAFANSEADRDYICDRYPNLMVYVDFOFHIMGLIPGIYIIS
YCIISIKLSHSHKQKRAKTKTVIILHFAFMLEYIYIGISIDSTILEITIQGEF
ENIVHKWISTTEALFHCOLNPLILYAFIGAKFKTSAOHLIVSRGSSLSLKGR
GHSVSSTSESSSFHSS"
intron 1053. .3184
/number=1
/gene="CXCR-4"
exon 3185. .4760
/number=2
/gene="CXCR-4"
3'UTR 4229. .4760
/gene="CXCR-4"
polyA_signal 4726. .4731
/gene="CXCR-4"
BASE COUNT 1160 a 1338 c 1287 g 1376 t
ORIGIN

Query Match 90.0%; Score 1564; DB 11; Length 5161;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ATATACACTTCAGATACTACACCGAGAAATGGGCTCAGGGGACTAGCTCCATGAAG 165
DB 1185 ATATACACTTCAGATACTACACCGAGAAATGGGCTCAGGGGACTAGCTCCATGAAG 3244
QY 156 GAACCGCTGTTCCGTAGAGAAATGCTAATTTCAATAAATCTTCTGCCACCACATCTAC 225
DB 3245 GAACCGCTGTTCCGTAGAGAAATGCTAATTTCAATAAATCTTCTGCCACCACATCTAC 3304
QY 226 TCCATCATCTTTTAAGTGCATTTGGGCATGATGGTGCATCTCGGTCATGGGTAC 285
DB 3305 TCCATCATCTTTTAAGTGCATTTGGGCATGATGGTGCATCTCGGTCATGGGTAC 3364
QY 286 CAGAGAACTGAGAGACATGACGACAGTACAGGCTGCACCTGTCACTGAGCCGACCTC 345
DB 3365 CAGAGAACTGAGAGACATGACGACAGTACAGGCTGCACCTGTCACTGAGCCGACCTC 3424
QY 346 CTCTTTGTCATCAGCGCTTCCTCTGGCGAGTTGATCCGTGGCAAACTGTGACTTGGG 405
DB 1185 CTCTTTGTCATCAGCGCTTCCTCTGGCGAGTTGATCCGTGGCAAACTGTGACTTGGG 405

DB 3425 CTCTTTGTCATCAGCGCTTCCTCTGGCGAGTTGATCCGTGGCAAACTGTGACTTGGG 3484
QY 406 AACTTCTTAAGCAAGGACATGCATGTATCTACACAGTCAACCTCTACAGCAGTGTCTC 465
DB 3485 AACTTCTTAAGCAAGGACATGCATGTATCTACACAGTCAACCTCTACAGCAGTGTCTC 3544
QY 466 ATCTGGCCTTCATCAGCTGAGGACCGGTACTGGCCATTCGTCACGCGCACCAACAGTCAG 525
DB 3545 ATCTGGCCTTCATCAGCTGAGGACCGGTACTGGCCATTCGTCACGCGCACCAACAGTCAG 3604
QY 526 AGGCCAAGAAAGCTGTGGCTGAAAAGTGGTCTATGTGGCGTCTGGATCCCTGCCCTC 585
DB 3605 AGGCCAAGAAAGCTGTGGCTGAAAAGTGGTCTATGTGGCGCTGTGGATCCCTGCCCTC 3664
QY 586 CTGCTGACTATCCGACCTTCATCTTTGCCAACGTCAGTGAAGCAGATGACAGATATATC 645
DB 3665 CTGCTGACTATCCGACCTTCATCTTTGCCAACGTCAGTGAAGCAGATGACAGATATATC 3724
QY 646 TGTACCGCTTCACCCCAATGACTGTGGGTGGTGTGTGCCAGTTTCAGACATCATG 705
DB 3725 TGTACCGCTTCACCCCAATGACTGTGGGTGGTGTGTGCCAGTTTCAGACATCATG 3784
QY 706 GTTGGCCTTATCTGCTGCTGATTTGTCATCCTGTCTGCTATTTGCATTAATCATCTCCAG 765
DB 3785 GTTGGCCTTATCTGCTGCTGATTTGTCATCCTGTCTGCTATTTGCATTAATCATCTCCAG 3844
QY 766 CTGTCACTATCCAAAGGCGCACAGAACCGCAAGGCCCTCAAGACCAAGATCTCTATC 825
DB 3845 CTGTCACTATCCAAAGGCGCACAGAACCGCAAGGCCCTCAAGACCAAGATCTCTATC 3904
QY 826 CTGGCTTCTTCGGCTGCTTGGCTGCTTACTACATTTGGATTCAGACATGACTCCTTCATC 885
DB 3905 CTGGCTTCTTCGGCTGCTTGGCTGCTTACTACATTTGGATTCAGACATGACTCCTTCATC 3964
QY 886 CTCTGGAATTCATCAAGCAAGGCTGTGAGTTGAGACACTGTGCACAGTGGATTTCC 945
DB 3965 CTCTGGAATTCATCAAGCAAGGCTGTGAGTTGAGACACTGTGCACAGTGGATTTCC 4024
QY 946 ATCACGAGGCGCTAGGCTTCTTCACATGTGTCTGTAACCCCATCTCTATGCTTCTT 1005
DB 4025 ATCACGAGGCGCTAGGCTTCTTCACATGTGTCTGTAACCCCATCTCTATGCTTCTT 4084
QY 1006 GGAAGCAATTTAAACCTCTGCCACAGCAGTCACTCTGTGAGCAGAGGCTCCAGC 1065
DB 4085 GGAAGCAATTTAAACCTCTGCCACAGCAGTCACTCTGTGAGCAGAGGCTCCAGC 4144
QY 1066 CTCAAGATCTCTCCAAAGGAAAGGAGGTGACATTTCACTGTTCCACTGAGTCTGAG 1125
DB 4145 CTCAAGATCTCTCCAAAGGAAAGGAGGTGACATTTCACTGTTCCACTGAGTCTGAG 4204
QY 1126 TCTTCAAGTTTCACTCAGCTAACAACAGATGTAAAGACTTTTATATAGATAATA 1185
DB 4205 TCTTCAAGTTTCACTCAGCTAACAACAGATGTAAAGACTTTTATATAGATAATA 4264
QY 1186 ACTTTTATTAAGTTACAACTTTTCAGATATAAAGACTGACCAATATGTACAGTTT 1245
DB 4265 ACTTTTATTAAGTTACAACTTTTCAGATATAAAGACTGACCAATATGTACAGTTT 4324
QY 1246 TATTGCTGTGGATTTTGTCTGTGTTCTTTAGTTTGTGAACTTAATTTGACTTA 1305
DB 4325 TATTGCTGTGGATTTTGTCTGTGTTCTTTAGTTTGTGAACTTAATTTGACTTA 4384
QY 1306 TTTATATAATTTTATTTTGTTCATATGTAGTGTCTGAGCAGAGCTGTGGCCAGT 1365
DB 4385 TTTATATAATTTTATTTTGTTCATATGTAGTGTCTGAGCAGAGCTGTGGCCAGT 4444
QY 1366 TCTTAGTGTGTATGTCTGTGTGAGAGCTGTAGAAAAGGAAGTGAACATTTCCAGAGC 1425
DB 4445 TCTTAGTGTGTATGTCTGTGTGAGAGCTGTAGAAAAGGAAGTGAACATTTCCAGAGC 4504
QY 1426 GTGTAGGAATCAAGTAAAGCTGAAATAGATCCCGAGCTGTTATAGCATAGATTAATCTCT 1485
DB 4505 GTGTAGGAATCAAGTAAAGCTGAAATAGATCCCGAGCTGTTATAGCATAGATTAATCTCT 4564

```
QY 1486 CCATCCCGTGAACGTTTTTCTGTTCTTAAGACGTGATTTGCTGTAGACATGGCAC 1545
|||||
DB 4565 CCATTCCTCGTGAACGTTTTTCTGTTCTTAAGACGTGATTTGCTGTAGACATGGCAC 4624
|||||
QY 1546 TTATACCAAGCCCAAGGTGTATAGAAATCGTGTTTTTCAGTTTTTCAGAGTGGGT 1605
|||||
DB 4625 TTATACCAAGCCCAAGGTGTATAGAAATCGTGTTTTTCAGTTTTTCAGAGTGGGT 4684
|||||
QY 1606 GATTTCAGCACTACAGTGTACGTCTGTATTAAGTTGTTAATAAAGTACATGTTAA 1665
|||||
DB 4685 GATTTCAGCACTACAGTGTACGTCTGTATTAAGTTGTTAATAAAGTACATGTTAA 4744
|||||
QY 1666 CTTA 1669
|||||
DB 4745 CTTA 4748

RESULT 8
HSCXCR4 3733 bp DNA PRI 02-APR-1998
LOCUS Homo sapiens CXCR4 gene.
DEFINITION Y14739
ACCESSION G3021393
NID Y14739.1 GI:3021393
VERSION Chemokine receptor; CXCR4 gene.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 3733)
AUTHORS Friedl, R., Moepf, B. and Gierschik, P.
TITLE Genomic organization and expression pattern of the human chemokine
receptor CXCR4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3733)
AUTHORS Friedl, R.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-1997) R. Friedl, Department of Pharmacology and
Toxicology, Albert-Einstein-Allee 11, University of Ulm, 89081 Ulm,
FRG
REMARK revised by submitter 31-MAR-1998
FEATURES
SOURCE Location/Qualifiers
1..3733
/organism="Homo sapiens"
/db_xref="taxon:9606"
/risue_type="blood"
join(<51..65,2198..>3241)
/gene="CXCR4"
<51..65
/gene="CXCR4"
/number=1
51..3241
/gene="CXCR4"
join(51..65,2198..3241)
/gene="CXCR4"
/codon_start=1
/product="chemokine receptor"
/protein_id="CAA75034.1"
/db_xref="PDB:e1285584"
/db_xref="PID:g3021394"
/translation="MEGISIYISDNYTEEMSGSDYSMEKPECFRENNANFNKIFLEPTI
YSIIETGIVNGELVLMGYOKLSMTDKYRLSLVADLLFVITLPPWADVAVNM
YGNPFCRAVHYVINYLSVILAFISIDRYLAIVHAINSORPKLLAEKVYVYG
WTPALLLTIPDIFANVSADDRITCDRYPNIMVYVPOHIMVGLIPGIVISC
YCIILSKSHSGHOKRAKTVIILILAFRCMLPYIIGISIDSLILEIKOGGEF
ENMTHKWIISTELALAFHCLNDILYFLGAKFKTISOHRLTIVSKGSLIKLSKGR
GGHSVSTSESSSFHSS"
66..2197
/gene="CXCR4"
/number=1
2198..>3241
intron
exon
```

```
BASE COUNT 822 a 955 c 980 g 976 t
ORIGIN
Query Match 88.4%; Score 1536; DB 10; Length 3733;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/number=2
QY 106 ATATACCTCAGTACTACACGAGGAATGGGCTCAGGGACATGATCTCATGAAG 165
|||||
DB 2198 ATATACCTCAGTACTACACGAGGAATGGGCTCAGGGACATGATCTCATGAAG 2257
|||||
QY 166 GAACCTGTTCCGTGAAGAAATGCTAATTTCAATAAATCTTCTGCCACATCTAC 225
|||||
DB 2258 GAACCTGTTCCGTGAAGAAATGCTAATTTCAATAAATCTTCTGCCACATCTAC 2317
|||||
QY 226 TCCATCATCTTTCTTAAGTGGCATTTGGGCAATGATTTGGTCACTCCGTCATGGTTAC 285
|||||
DB 2318 TCCATCATCTTTCTTAAGTGGCATTTGGGCAATGATTTGGTCACTCCGTCATGGTTAC 2377
|||||
QY 286 CAGAAGAACTGAGAGATGACGACAGTACAGCTGTCAGTGGCCGACCTC 345
|||||
DB 2378 CAGAAGAACTGAGAGATGACGACAGTACAGCTGTCAGTGGCCGACCTC 2437
|||||
QY 346 CTCTTTGTCATCAGGCTTCCCTTCTGGGCACTGATGCGGTGGCAACTGTTGGG 405
|||||
DB 2438 CTCTTTGTCATCAGGCTTCCCTTCTGGGCACTGATGCGGTGGCAACTGTTGGG 2497
|||||
QY 406 AACTTCCATGAGAGGCGATCATGTATCTACACAGTCAACCTCTACAGAGTGTCTC 465
|||||
DB 2498 AACTTCCATGAGAGGCGATCATGTATCTACACAGTCAACCTCTACAGAGTGTCTC 2557
|||||
QY 466 ATCTGGCCTTCATCATGCTGTGACCGCTACCTGGCCATGTCACAGCCACCAACAGTAG 525
|||||
DB 2558 ATCTGGCCTTCATCATGCTGTGACCGCTACCTGGCCATGTCACAGCCACCAACAGTAG 2617
|||||
QY 526 AGGCCAAGAACTGTTGGCTGAAAAAGTGTTATGTTGGCGTGTGATCCCTGCCCTC 585
|||||
DB 2618 AGGCCAAGAACTGTTGGCTGAAAAAGTGTTATGTTGGCGTGTGATCCCTGCCCTC 2677
|||||
QY 586 CTGCTGACATATCCGACCTTCACTCTTCCCAACGTCAGTGGAGGACAGATATATATC 645
|||||
DB 2678 CTGCTGACATATCCGACCTTCACTCTTCCCAACGTCAGTGGAGGACAGATATATATC 2737
|||||
QY 646 TGTGACCGCTTCTACCCCAATGACTTGTGGGTGTTGTGTTCCATGTTTCAGCACTATATG 705
|||||
DB 2738 TGTGACCGCTTCTACCCCAATGACTTGTGGGTGTTGTGTTCCATGTTTCAGCACTATATG 2797
|||||
QY 706 GTTGGCCTTATCTCGCTGCTGATTTGTCATCTGCTGTGATTTGATTTATCATCTCCAAAG 765
|||||
DB 2798 GTTGGCCTTATCTCGCTGCTGATTTGTCATCTGCTGTGATTTGATTTATCATCTCCAAAG 2857
|||||
QY 766 CTGTCACTTCCAAAGGCGCCAGAGCGCAAGGCCCTCAAGACACAGTCAATCCTCATC 825
|||||
DB 2858 CTGTCACTTCCAAAGGCGCCAGAGCGCCCTCAAGACACAGTCAATCCTCATC 2917
|||||
QY 826 CTGGCTTTTCTGCTGTTGGCTGCTTACTACTATTTGGATCAGATGAGTCACTCTCATC 885
|||||
DB 2918 CTGGCTTTTCTGCTGTTGGCTGCTTACTACTATTTGGATCAGATGAGTCACTCTCATC 2977
|||||
QY 886 CTCTCGGAATATCAAGGAAGGTTGATTTGAGAACTGTGCAAGTGAATTTCC 945
|||||
DB 2978 CTCTCGGAATATCAAGGAAGGTTGATTTGAGAACTGTGCAAGTGAATTTCC 3037
|||||
QY 946 ATACCGAGGCGCTTACTTCTTCCACTGTGTCTGAACCCATCCTCTATGCTTTCTT 1005
|||||
DB 3038 ATACCGAGGCGCTTACTTCTTCCACTGTGTCTGAACCCATCCTCTATGCTTTCTT 3097
|||||
QY 1006 GAGCCAAATTTAAACCTCTGCGCAGACGCACTACCTCTGTGAGAGAGGTTCCAG 1065
|||||
DB 3098 GAGCCAAATTTAAACCTCTGCGCAGACGCACTACCTCTGTGAGAGAGGTTCCAG 3157
|||||
```

OY	1065	CTCAAGATCCCTCTCCAAAGCAAGCCAGGCGACATTCATCTGTTCCACGTAGCTAG	11259
Db	3158	CTCAAGATCCCTCTCCAAAGCAAGCCAGGCGACATTCATCTGTTCCACGTAGCTAG	32127
OY	1126	TCTTCAAGTTTCACTCCAGCTAACACAGATGTAAAGACTTTTTTTATACGATAAATA	11859
Db	3218	TCTTCAAGTTTCACTCCAGCTAACACAGATGTAAAGACTTTTTTTATACGATAAATA	32777
OY	1185	ACTTTTATTAAGTACACATTTTTCAGATATATAAAGACTGACCATTATTGACAGTTT	12455
Db	3278	ACTTTTATTAAGTACACATTTTTCAGATATATAAAGACTGACCATTATTGACAGTTT	33377
OY	1245	TATTCCTTGTTGGATTTTGTGCTGTGTTCTTTAGTTTGTGGAAGTAAATACACTTA	13059
Db	3338	TATTCCTTGTTGGATTTTGTGCTGTGTTCTTTAGTTTGTGGAAGTAAATACACTTA	33977
OY	1306	TTTATATAAATTTTTTTTGTTCATATTGATGTGTCTACGACAGACCTGTGGCCAAGT	13655
Db	3398	TTTATATAAATTTTTTTTGTTCATATTGATGTGTCTACGACAGACCTGTGGCCAAGT	34577
OY	1356	TCTTAGTTGCTATATGCTCTGTGGTAGAGACTGTAGAAAAGGAACTGAACATTCCAGAGC	14255
Db	3458	TCTTAGTTGCTATATGCTCTGTGGTAGAGACTGTAGAAAAGGAACTGAACATTCCAGAGC	35177
OY	1425	GTTAGTAGCATACGTAACCTGAAGAAATATATCCCGAGCTGTTATGACATAGATATCTCT	14855
Db	3518	GTTAGTAGCATACGTAACCTGAAGAAATATATCCCGAGCTGTTATGACATAGATATCTCT	35777
OY	1485	CCATTCCTCCGTGGAACGTTTTCTGTGTTCTTAACACGTATTTGCTGTAGAAATGAGCAC	15455
Db	3578	CCATTCCTCCGTGGAACGTTTTCTGTGTTCTTAACACGTATTTGCTGTAGAAATGAGCAC	36377
OY	1545	TTATTAACCAAGCCCAAAAGTGTGTATAGAAATGCTGTTTTTCAGTTTTCAGAGATGGGTT	16055
Db	3638	TTATTAACCAAGCCCAAAAGTGTGTATAGAAATGCTGTTTTTCAGTTTTCAGAGATGGGTT	36977
OY	1605	GATTCAGCACTACAGGTGACGCTTGATTTAAG	1641
Db	3698	GATTCAGCACTACAGGTGACGCTTGATTTAAG	3733

```

QY 466 ATCTGGCCCTCATAGTGTGAGACCGCTACCTGGCCATGCTCCACGCCACCAACAGTCAAG 525
    |||
Db 5228 ATCCCTGGCCCTCATAGTGTGAGACCGCTACCTGGCCATGCTCCACGCCACCAACAGTCAAG 5287
QY 526 AGCCCAAGAGAGCTGTGGCTGAAAAGGSGTGTATGTGGGCTGTGAGATCCCTGCCCTC 585
    |||
Db 5288 AGCCCAAGAGAGCTGTGGCTGAAAAGGSGTGTATGTGGGCTGTGAGATCCCTGCCCTC 5347
QY 586 CTGCTGACTATTCCCGACCTTCATCTTTGCCACAGTCAGTGGCAGATGACATATATATC 645
    |||
Db 5348 CTGCTGACTATTCCCGACCTTCATCTTTGCCACAGTCAGTGGCAGATGACATATATATC 5407
QY 646 TGTGACCGCTTCAACCCCAATGACTGTGGGTGGTGTGTCAGTTTCCAGCATCATG 705
    |||
Db 5408 TGTGACCGCTTCAACCCCAATGACTGTGGGTGGTGTGTCAGTTTCCAGCATCATG 5467
QY 706 GTTGGGCTTATCCGCTGTGATGTGATGCTGCTGCTATTTGACATTAATCATCTCCAG 765
    |||
Db 5458 GTTGGGCTTATCCGCTGTGATGTGATGCTGCTGCTATTTGACATTAATCATCTCCAG 5527
QY 766 CTGTCACTCTCAAGGGCCACAGAGCGCAAGGCCCTCAAGACACAGTCATCTCATC 825
    |||
Db 5528 CTGTCACTCTCAAGGGCCACAGAGCGCAAGGCCCTCAAGACACAGTCATCTCATC 5587
QY 826 CTGCTTTCTTCCGCTGTGGCTGCTTACATGATGGAGTGCATGCACTCCTTCATC 885
    |||
Db 5588 CTGCTTTCTTCCGCTGTGGCTGCTTACATGATGGAGTGCATGCACTCCTTCATC 5647
QY 886 CTCTGGAATATCATCAAGCAAGGGTGTGAGTTTGAACACATGTGCACAGTGTATTC 945
    |||
Db 5648 CTCTGGAATATCATCAAGCAAGGGTGTGAGTTTGAACACATGTGCACAGTGTATTC 5707
QY 946 ATCACCAGAGCCCTAGCTTTTCTTCACCTGTGTGAAACCCATCCTATGCTTCTT 1005
    |||
Db 5708 ATCACCAGAGCCCTAGCTTTTCTTCACCTGTGTGAAACCCATCCTATGCTTCTT 5767
QY 1006 GGACCCAAATTTAAACCTGTGCCAGCAGCAGTCTGTGAGAGAGGGTCCAGC 1065
    |||
Db 5768 GGACCCAAATTTAAACCTGTGCCAGCAGCAGTCTGTGAGAGAGGGTCCAGC 5827
QY 1066 CTCAAGATCTCTCCAAAGGAAAGGAGGTGAGACATTAATCTGTTCCACGTGAGTCAAG 1125
    |||
Db 5828 CTCAAGATCTCTCCAAAGGAAAGGAGGTGAGACATTAATCTGTTCCACGTGAGTCAAG 5887
QY 1126 TCTTCAATTTTCACTCCAGTCAACACAGATGTAAAGACTTTTATACGATTAATA 1185
    |||
Db 5888 TCTTCAATTTTCACTCCAGTCAACACAGATGTAAAGACTTTTATACGATTAATA 5947
QY 1186 ACTTTTTTAAAGTACATTTTTCAGATATTAATAAGCTGACCAATATTGACGTTT 1245
    |||
Db 5948 ACTTTTTTAAAGTACATTTTTCAGATATTAATAAGCTGACCAATATTGACGTTT 6007
QY 1246 TATTCCTGTGGATTTTGTCTTGTGTTCTTATGTTTCTGAAAGTTAATGACTTA 1305
    |||
Db 6008 TATTCCTGTGGATTTTGTCTTGTGTTCTTATGTTTCTGAAAGTTAATGACTTA 6067
QY 1306 TTTATATAATTTTGTGTTTCTTATATGATGTGTGTAGCAGAGACCTGTGGCAAGT 1365
    |||
Db 6068 TTTATATAATTTTGTGTTTCTTATATGATGTGTGTAGCAGAGACCTGTGGCAAGT 6127
QY 1366 TCTTAGTTCCTGATGTCTGCTGTGAGAGCTGAGAAAAGGAACTGAACATTTCCAGAGC 1425
    |||
Db 6128 TCTTAGTTCCTGATGTCTGCTGTGAGAGCTGAGAAAAGGAACTGAACATTTCCAGAGC 6187
QY 1426 GTGTAGTGAATCAAGTAAGCTGAGAAATGATATCCCAAGCTGTTATGATATCATCTCT 1485
    |||
Db 6188 GTGTAGTGAATCAAGTAAGCTGAGAAATGATATCCCAAGCTGTTATGATATCATCTCT 6247
QY 1486 CCATTCCTCGTGAAGCTTTTCTGTTCTTAAGAGCTGATTTTCTGTAGAAAGATGGCAC 1545
    |||
Db 6248 CCATTCCTCGTGAAGCTTTTCTGTTCTTAAGAGCTGATTTTCTGTAGAAAGATGGCAC 6307
QY 1546 TTATAACCAAAAGCCAAAGTGTGATAGAAATGCTGTTTTCAGTTTTCAGGAGTGGGT 1605

```

```

Db 6308 TTATAACCAAAAGCCAAAGTGTGATAGAAATGCTGTTTTCAGTTTTCAGGAGTGGGT 6367
QY 1606 GATTTCAGACCTACAGTGTGATAGTGTATTAAGTTTAATAAGTACATGTAA 1665
    |||
Db 6368 GATTTCAGACCTACAGTGTGATAGTGTATTAAGTTTATTAAGTACATGTAA 6427
QY 1666 CTTA 1669
    |||
Db 6428 CTTA 6431

RESULT 10
AF052572 8747 bp DNA PRI 07-SEP-1998
LOCUS Homo sapiens chemokine receptor CXCR4 gene, promoter region and
DEFINITION complete cds.
ACCESSION AF052572
NID 93549254
VERSION AF052572.1 GI:3549254
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 8747)
            Caruz,A., Samsom,M., Alonso,J.M., Alcamí,J., Balleux,F.,
            Virelizier,J.L., Parmentier,M. and Arenzana-Seisdedos,F.
            Genomic organization and promoter characterization of human CXCR4
            gene.
            FEBS Lett. 426 (2), 271-278 (1998)
TITLE JOURNAL MEDLINE
        96258970
REFERENCE 2 (bases 1 to 8747)
            Caruz,A., Samsom,M., Virelizier,J.L., Parmentier,M. and
            Arenzana-Seisdedos,F.
            Direct Submission
TITLE JOURNAL
        Submitted (06-MAR-1998) Immunologie Virale, Institut Pasteur, 28
        Rue Dr Roux, Paris 75724, France
FEATURES
        source
            1..8747
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="24D3"
            /note="BAC clone from Research Genetics"
            1..2642
            /join(2643..2735,4868..6429)
            /product="chemokine receptor CXCR4"
            /join(2721..2735,4868..5911)
            /codon_start=1
            /product="chemokine receptor CXCR4"
            /protein_id="AAC34581.1"
            /db_xref="PIR:g3549255"
            /db_xref="PIR:g3549255"
            /translation="MEGISITSDNYTEMGSGDYDKMKPEPRENANFNKFLPTI
            YSILFLGIVNGVLIVMGYOKRLRSMTRKYLRLHSVADLLFVITLPEFAVDVAVNM
            YFGEFLKRAVAVITVNLVYSVLILAFISIDRYLAIVAHNSQPRRLAEKVAVG
            WIPALLILDEIPFANVSEADDRYICDRFENDLAVVFOHIMVLIIPGIVLSC
            YCIIISLTSKSGOKRKALKTKTYIILAFACAPLYIGISDSFLLIEKGGCEP
            ENYHKMISTIEALFHCINPLIYVFLAKFKTSNQHALTSVRSRSSLIKLSKGR
            GGHSSVSSESSSFHSS"
BASE COUNT 2161 a 2173 c 1966 g 2447 t
ORIGIN
Query Match 87.1%; Score 1513; DB 11; Length 8747;
Best Local Similarity 99.9%; Fred. No. 0; Mismatches 1; Indels 0; Gaps 0;
Matches 1563; Conservative 0;
QY 106 ATATACCTTCAGATACACGAGAAATGGCTCAGGAGCTATGACTCATGAAG 165
    |||
Db 4868 ATATACCTTCAGATACACGAGAAATGGCTCAGGAGCTATGACTCATGAAG 4927
QY 166 GAACCTGTTCCGTGAGAAAGAAATGCTAAATTCAATAAATCTTCTGCCACATCTAC 225

```

```
|||||
Db 4928 GAACCCGTTCCGTGAGAAATGCTAATTTCATATAAACTCTCCGCCACCACATCTAC 4987
QY 226 TCCATCATCTTTAACTGGCATTTGGGCAATGATGGTCACTCCTGGTCATAGGTTAC 285
Db 4988 TCCATCATCTTTAACTGGCATTTGGGCAATGATGGTCACTCCTGGTCATAGGTTAC 5047
QY 286 CAGAAGAACTGAGAGCATGACGAGCAAGTACAGGCTGACCTGTCAGTGGCGAGCTC 345
Db 5048 CAGAAGAACTGAGAGCATGACGAGCAAGTACAGGCTGACCTGTCAGTGGCGAGCTC 5107
QY 346 CTCCTTGTGATCAGCGCTTCCCTTCTGGGCAATGATGGTCACTGCGCAAACTGGTACTTTGG 405
Db 5108 CTCCTTGTGATCAGCGCTTCCCTTCTGGGCAATGATGGTCACTGCGCAAACTGGTACTTTGG 5167
QY 406 AACTTCCATGCAAGGCAATGCTGATCTACAGTCAAGCTCTACAGCAATGCTC 465
Db 5168 AACTTCCATGCAAGGCAATGCTGATCTACAGTCAAGCTCTACAGCAATGCTC 5227
QY 466 ATCTGCGCTTGCATCAGTGTGACCGCTACCTGGCCATGCTCCAGCCCAACAGTACAG 525
Db 5228 ATCTGCGCTTGCATCAGTGTGACCGCTACCTGGCCATGCTCCAGCCCAACAGTACAG 5287
QY 526 AGGCCAAGCAAGCTGTGGCTGAAGAGTGTGTATGTTGGCGTCTGGATCCTGCCCTC 585
Db 5288 AGGCCAAGCAAGCTGTGGCTGAAGAGTGTGTATGTTGGCGTCTGGATCCTGCCCTC 5347
QY 586 CTGCTGACTATTCGCGACTTCATCTTGGCAAGCTGAGGAGGCAATGACATATATC 645
Db 5348 CTGCTGACTATTCGCGACTTCATCTTGGCAAGCTGAGGAGGCAATGACATATATC 5407
QY 646 TGTGACCGCTTGTACCCCAATGACTTGTGGTGGTGTGTGTCCAGTTTACGACATCATG 705
Db 5408 TGTGACCGCTTGTACCCCAATGACTTGTGGTGGTGTGTGTCCAGTTTACGACATCATG 5467
QY 706 GTTGGCTTATCTGCTGCTGTGTATGTATCTCTGCTGCTATGATATCATCTCCAG 765
Db 5468 GTTGGCTTATCTGCTGCTGTGTATGTATCTCTGCTGCTATGATATCATCTCCAG 5527
QY 766 CTGTCACTTCCAGAGGCGCCAGAGGCGCCCAAGGCGCTCAAGCCAGTCACTCTCATC 825
Db 5528 CTGTCACTTCCAGAGGCGCCAGAGGCGCCCAAGGCGCTCAAGCCAGTCACTCTCATC 5587
QY 826 CTGGCTTTCTTGCGCTGTGTGGCTGCTTACTACATTTGGATGAGATCGACTCTCTCATC 885
Db 5588 CTGGCTTTCTTGCGCTGTGTGGCTGCTTACTACATTTGGATGAGATCGACTCTCTCATC 5647
QY 886 CTCTCGAATATCATCAGAGGCGTGTGATTTGAGAACACTGTGCAAGTGGATTTC 945
Db 5648 CTCTCGAATATCATCAGAGGCGTGTGATTTGAGAACACTGTGCAAGTGGATTTC 5707
QY 946 ATCAGCGAGCGCTAGCTTTCTTCCAGTGTGTCTGTAACCCCATCTCTATGCTTCTT 1005
Db 5708 ATCAGCGAGCGCTAGCTTTCTTCCAGTGTGTCTGTAACCCCATCTCTATGCTTCTT 5767
QY 1006 GGAGCCAAATTTAAACCTCTGCGCCAGCAGCACTACTCTGTGAGAGGGTCCAGC 1065
Db 5768 GGAGCCAAATTTAAACCTCTGCGCCAGCAGCACTACTCTGTGAGAGGGTCCAGC 5827
QY 1066 CTCAGAGTCTCTCCAAAGGAAGGAGGTGGACATTCATCTGTTCCACTGAGTCTGAG 1125
Db 5828 CTCAGAGTCTCTCCAAAGGAAGGAGGTGGACATTCATCTGTTCCACTGAGTCTGAG 5887
QY 1126 TCTTCAAGTTTACTCAGCTAACACAGATGTAAAGACTTTTTTTATAGCATTAATA 1185
Db 5888 TCTTCAAGTTTACTCAGCTAACACAGATGTAAAGACTTTTTTTATAGCATTAATA 5947
QY 1186 ACTTTTGTAGTACATTTTGTAGATATATAAGACTGACCATATTTGTACAGTTT 1245
Db 5948 ACTTTTGTAGTACATTTTGTAGATATATAAGACTGACCATATTTGTACAGTTT 6007
QY 1246 TATTGCTTGTGATTTTGTGTGTCTTAACTTTTGTGAGTTAATGACTTA 1305
|||||
```

```
Db 6008 TATTGCTTGTGATTTTGTGTGTCTTAACTTTTGTGAGTTAATGACTTA 6067
QY 1306 TTTATATAAATTTTGTGTGTATTTGATGTGTGTCTAGCGAGCACTGTGGCCAAGT 1365
Db 6068 TTTATATAAATTTTGTGTGTATTTGATGTGTGTCTAGCGAGCACTGTGGCCAAGT 6127
QY 1366 TCTTGTAGTGTATGTCTGCTGTAGGACTGTAGAAAAGGAGCACTGAACATTCAGAGC 1425
Db 6128 TCTTGTAGTGTATGTCTGCTGTAGGACTGTAGAAAAGGAGCACTGAACATTCAGAGC 6187
QY 1426 GTGTAGTGAATCAGTGAAGCTAGAAAATGATCCCGAGCTGTATATCATAGATATCTGT 1485
Db 6188 GTGTAGTGAATCAGTGAAGCTAGAAAATGATCCCGAGCTGTATATCATAGATATCTGT 6247
QY 1486 CCAATCCCGTGGAGCACTTTTCTGTTTAAAGCTGATTTTGTGTAGAAATGGCAC 1545
Db 6248 CCAATCCCGTGGAGCACTTTTCTGTTTAAAGCTGATTTTGTGTAGAAATGGCAC 6307
QY 1546 TTATAACCAAGGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTT 1605
Db 6308 TTATAACCAAGGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTT 6367
QY 1606 GATTTCAGCACTACAGTGTACAGTCTTGTATTAGCTTTAATAAAGTACATGTTAA 1665
Db 6368 GATTTCAGCACTACAGTGTACAGTCTTGTATTAGCTTTAATAAAGTACATGTTAA 6427
QY 1666 CTTA 1669
Db 6428 CTTA 6431

RESULT 11
HUMSTR 1637 bp mRNA PRI 10-MAR-1994
LOCUS Homo sapiens seven transmembrane segment receptor mRNA, complete
DEFINITION
ACCESSION M99293
NID 9292516
VERSION M99293.1 GI:292516
KEYWORDS seven transmembrane segment receptor.
SOURCE Homo sapiens (human).
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1637)
Fedorov, B., Melhado, I.G., Duncan, A.M., Delaney, A.,
Schappert, K., Clark-Lewis, I., and Jirik, F.R.
Molecular cloning of the cDNA and chromosomal localization of the
gene for a putative seven-transmembrane segment (7-TMS) receptor
isolated from human spleen
JOURNAL Genomics 16, 707-712 (1993)
MEDLINE 93315164
FEATURES
source
misc_feature
1..1637
location=Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="2q21"
1
/feature="GDB identifier D25201E; S9a localized by in situ
hybridization to human chromosome 2q21"
85..1143
/note="93% homology to bovine neuropeptide Y3; precursor;
putative"
/codon_start=1
/product="seven transmembrane segment receptor"
/protein_id="AA16617.1"
/db_xref="PDB:9292517"
/db_xref="GI:292517"
/translation="MEGISITVSDNYEEMSGSDYSKPECFREENANFKIFLPTI
YSILFLGIVNGLVILVMGYOKLRMTDKYRLHLVADLLFVITLPMFVADVAVNM
YFNFELCAVHIVYIVNLVSYVLLAFISLDRIYLAIVHATNSQRPRLLAERVYGV
WIPALLETIPFIFANVSEADDRICDRFYNDLWVYVFOROHIMVGLILPGIYLISC
KIIISKSHSKGKRAKLTIVILLIAFACWLPYIGISIDSLLEILIKOGER
```


ENTVHKMISTEALAFPHCLNPILYAFILGAKFKTSQHALTSVSRGSLKILSKGR
 GGHSSVTESSSFHSS"
 BASE COUNT 394 a 399 c 363 g 481 t
 ORIGIN

Query Match 73.6% Score 1278; DB 9; Length 1637;
 Best Local Similarity 100.0% Pred. No. 0;
 Matches 1278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 GCAGAGTACCCGAGGAGGCTGAGTGTCTCCAGTAGCCACCGCATCTGGAGAACGAGCGT 86
 Db 21 GCAAGTACGAGGAGGCTGAGTGTCTCCAGTAGCCACCGCATCTGGAGAACGAGCGT 80
 QY 87 TACCATGAGGAGGATCATATATACACTTCAGATACACTACACGAGAAATGGGCTCAGG 146
 Db 81 TACCATGAGGAGGATCATATATACACTTCAGATACACTACACGAGAAATGGGCTCAGG 140
 QY 147 GGAATATGACTCCATGAGAGAACCCCTGTTCCGTGAGAGAAATGCTAATTTCAATAAAT 206
 Db 141 GGAATATGACTCCATGAGAGAACCCCTGTTCCGTGAGAGAAATGCTAATTTCAATAAAT 200
 QY 207 CTTCCTGCCCCACCATCTACCTCATCATCTTTCTTAACCTGGCATTTGGGCAATGGATTGGT 266
 Db 201 CTTCCTGCCCCACCATCTACCTCATCATCTTTCTTAACCTGGCATTTGGGCAATGGATTGGT 260
 QY 267 CATCTGTGTCATGGGTTTCCAGAGAAACTGAGAGAACGAGGAGGATGACAGGCTGCA 326
 Db 261 CATCTGTGTCATGGGTTTCCAGAGAAACTGAGAGAACGAGGAGGATGACAGGCTGCA 320
 QY 327 CCTGTGATGAGCCGACCTCTCTTTGTTCATGACAGCTTCCCTTCTGGGAGATTGATGCCGT 386
 Db 321 CCTGTGATGAGCCGACCTCTCTTTGTTCATGACAGCTTCCCTTCTGGGAGATTGATGCCGT 380
 QY 387 GGCAGAACTGTACTTTGGGAATTCCTATGCAAGGCACTTCATGTCATCTACACAGTAA 446
 Db 381 GGCAGAACTGTACTTTGGGAATTCCTATGCAAGGCACTTCATGTCATCTACACAGTAA 440
 QY 447 CCTGTACAGAGTGTCTTCATCTGAGCCCTTCATGTCAGACCCCTACCTGGCCATCGT 506
 Db 441 CCTGTACAGAGTGTCTTCATCTGAGCCCTTCATGTCAGACCCCTACCTGGCCATCGT 500
 QY 507 CCAGGCCACCAACAGCTGAGAGGCCCAAGAGCTGTGGCTGAAAAAGTGCTATATGG 566
 Db 501 CCAGGCCACCAACAGCTGAGAGGCCCAAGAGCTGTGGCTGAAAAAGTGCTATATGG 560
 QY 567 CGTGTGATCCCTGCCCCCTCTCTGCTGACTATTTCCGACATTCATCTTTGCCAAGCTAGGA 626
 Db 561 CGTGTGATCCCTGCCCCCTCTCTGCTGACTATTTCCGACATTCATCTTTGCCAAGCTAGGA 620
 QY 627 GGCAGATGACAGATATATCTGTGACGGCTTCTACCCCAATGACTTGTGGGTGGTGT 686
 Db 621 GGCAGATGACAGATATATCTGTGACGGCTTCTACCCCAATGACTTGTGGGTGGTGT 680
 QY 687 CCAGTTTACGACATCATGTTGGCTTATCTGCTGCTGATTTGCTATCTCTCTGCTA 746
 Db 681 CCAGTTTACGACATCATGTTGGCTTATCTGCTGCTGATTTGCTATCTCTCTGCTA 740
 QY 747 TTGCAATATATCTGCAAGCTGTACACTTCGAAGGCCACCGAAGGCCAAGGCCCTCAA 806
 Db 741 TTGCAATATATCTGCAAGCTGTACACTTCGAAGGCCACCGAAGGCCAAGGCCCTCAA 800
 QY 807 GAGCCAGAGTACCTCATCTGGCTTTCTTGCCCTTTGGGCTTGTACTACATTTGGAT 866
 Db 801 GAGCCAGAGTACCTCATCTGGCTTTCTTGCCCTTTGGGCTTGTACTACATTTGGAT 860
 QY 867 CAGCATCGAATCTCTCATCTCTCTGGAATTCATCAAGCAAGGGTGTGAGTTTGAGAACAC 926
 Db 861 CAGCATCGAATCTCTCATCTCTCTGGAATTCATCAAGCAAGGGTGTGAGTTTGAGAACAC 920
 QY 927 TGTGACAAAGTGGATTCATACGAGAGGCCCTAGCTTTCTTCCACTGTGTGTGAACCC 986
 Db 921 TGTGACAAAGTGGATTCATACGAGAGGCCCTAGCTTTCTTCCACTGTGTGTGAACCC 980

QY 987 CATTCTCTATGCTTCTTCCTTGGAGCCAAATTTAAACCTCTGCCAGACGACTCACTC 1046
 Db 981 CATTCTCTATGCTTCTTCCTTGGAGCCAAATTTAAACCTCTGCCAGACGACTCACTC 1040
 QY 1047 TGTGAGCAGAGGGTCCAGCTCAGATCCCTCTCCAAAGAAACGAGGTGGACATTCATC 1106
 Db 1041 TGTGAGCAGAGGGTCCAGCTCAGATCCCTCTCCAAAGAAACGAGGTGGACATTCATC 1100
 QY 1107 TGTTCACAGTGTGAGTGTCTTCAAGTTTCTACCTCCAGTAAACAGATGATAAAGACT 1166
 Db 1101 TGTTCACAGTGTGAGTGTCTTCAAGTTTCTACCTCCAGTAAACAGATGATAAAGACT 1160
 QY 1167 TTTTATATAGATAATAAATCTTTTATAGTTACATTTTTCAGATATAAAGACTG 1226
 Db 1161 TTTTATATAGATAATAAATCTTTTATAGTTACATTTTTCAGATATAAAGACTG 1220
 QY 1227 ACCAATATGTGACAGTTTATTTGCTGTTGGATTTTGTCTGTGTTCTTTAGTTT 1286
 Db 1221 ACCAATATGTGACAGTTTATTTGCTGTTGGATTTTGTCTGTGTTCTTTAGTTT 1280
 QY 1287 GTGAAGTTTAATTGACTT 1304
 Db 1281 GTGAAGTTTAATTGACTT 1298

RESULT 12

HUMGPCR 1670 bp mRNA PRI 10-NOV-1993
 LOCUS Human (clone L5) orphan G protein-coupled receptor mRNA, complete
 DEFINITION

ACCESSION L06797
 MID L06797.1 GI:414929
 VERSION 1
 KEYWORDS G protein-coupled receptor; neuropeptide receptor.
 SOURCE Human adult lung cDNA to mRNA.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 1670)
 Herzog, H., Hort, Y.J., Shine, J. and Seblie, L.A.
 Molecular cloning, characterization, and localization of the human
 homolog to the reported bovine NPY Y3 receptor: lack of NPY binding
 and activation

TITLE

homolog to the reported bovine NPY Y3 receptor: lack of NPY binding
 and activation

JOURNAL

DNA Cell Biol. 12, 465-471 (1993)
 MEDLINE 93319629
 COMMENT On Nov 10, 1993 this sequence version replaced gi:414929.

FEATURES

Location/Qualifiers
 1..1670
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /dev_stage="adult"
 /tissue_type="lung"
 /map="2"

CDS

73..1131
 /codon_start=1
 /product="G protein-coupled receptor"
 /protein_id="AA03209.1"
 /db_xref="PID:g414928"
 /translation="MEGISITYSDNTYEMSGSDYDSMKPECFEENANFNKLEPLTI
 YSIFLUGIYNGVILVMGOKRLREMTDKYRLHLSVADLLPITLTPFMAVDVAVVY
 YFGEFLKRAHVITYVNLVYSSVLLLATISDRYLAITYHATRSQPRLLAEKVVVYV
 WIPALLITIDPFIPIANVSEADRYICDRFPNDLVVDFQOHIMVGLIGVILISC
 YCIIISKLSHGKHKRAKLTIVILILAEFAWLPYIGISIDSFILLEIKGCE
 ENTVHKMISTEALAFPHCLNPILYAFILGAKFKTSQHALTSVSRGSLKILSKGR
 GGHSSVTESSSFHSS"

polyA_site
 BASE COUNT 418 a 400 c 361 g 491 t
 ORIGIN

Query Match

71.7% Score 1245; DB 9; Length 1670;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 1655; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

```

QY 30 AAGTGAAGGAGGAGGCTGAGTGTCTCCAGTAGCCACCGCATCTGGAGAACCCAGCGGTTAC 89
    |||
Db 12 AAGTGAAGGAGGAGGCTGAGTGTCTCCAGTAGCCACCGCATCTGGAGAACCCAGCGGTTAC 71
    |||
QY 90 CATGGAAGGAGATGATATACATTCAGATTAATACACGAGGAATGGGCTAGGGGA 149
    |||
Db 72 CATGGAAGGAGATGATATACATTCAGATTAATACACGAGGAATGGGCTAGGGGA 131
    |||
QY 150 CTATGATCCATGAAGAACCTGTTTCGCTGAAGAAATGCTAATTTCATAAATCTT 209
    |||
Db 132 CTATGATCCATGAAGAACCTGTTTCGCTGAAGAAATGCTAATTTCATAAATCTT 191
    |||
QY 210 CCGGCCACCATCTACTCCATCATCTTCTTAACCTGGCATGTGGGCAATGGATTGTCAT 269
    |||
Db 192 CCGGCCACCATCTACTCCATCATCTTCTTAACCTGGCATGTGGGCAATGGATTGTCAT 251
    |||
QY 270 CCTGTCATGGGTTACGAGAAGAACTGAGAGCATAGCGACAGTACAGGCTGCACCT 329
    |||
Db 252 CCGTCATGGGTTACGAGAAGAACTGAGAGCATAGCGACAGTACAGGCTGCACCT 311
    |||
QY 330 GTCATGAGCGGACCTCTCTTTTTCATCAGCGTCCCTTCTGGGCAAGTTGATGCCGTGGC 389
    |||
Db 312 GTCATGAGCGGACCTCTCTTTTTCATCAGCGTCCCTTCTGGGCAAGTTGATGCCGTGGC 371
    |||
QY 390 AACTGTGATCTTGGGAGACTTCATGCAAGGAGATGTCATGTCATCTACAGCACT 449
    |||
Db 372 AACTGTGATCTTGGGAGACTTCATGCAAGGAGATGTCATGTCATCTACAGCACT 431
    |||
QY 450 CTACAGCAGTGTCTCATCTCTGACCTTCATCAGTCTGGACCGCTACCTGGCCATCGTCCA 509
    |||
Db 432 CTACAGCAGTGTCTCATCTCTGACCTTCATCAGTCTGGACCGCTACCTGGCCATCGTCCA 491
    |||
QY 510 CGGCACCAACAGTCAGAGGCCAAGGAAGCTGTGGCTGAAGAAAGTGCTATGTGGCGT 569
    |||
Db 492 CGGCACCAACAGTCAGAGGCCAAGGAAGCTGTGGCTGAAGAAAGTGCTATGTGGCGT 551
    |||
QY 570 CTGATCCCTGCGCTCTGCTGACATATCCGAGCTTCATCTTGGCCAGCTCAGTGAAGC 629
    |||
Db 552 CTGATCCCTGCGCTCTGCTGACATATCCGAGCTTCATCTTGGCCAGCTCAGTGAAGC 611
    |||
QY 630 AGATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGGTGTGTGCCA 689
    |||
Db 612 AGATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGGTGTGTGCCA 671
    |||
QY 690 GTTTCAGCAGATCATGCTTGGCTTATCCTGCTGTGATTTGTATCTGCTCGCTATTTG 749
    |||
Db 672 GTTTCAGCAGATCATGCTTGGCTTATCCTGCTGTGATTTGTATCTGCTCGCTATTTG 731
    |||
QY 750 CATTATCATCTCAAGCTGTACACTCTCAAGAGGCCACAGAACGCGAAGGCCCTCAAGAC 809
    |||
Db 732 CATTATCATCTCAAGCTGTACACTCTCAAGAGGCCACAGAACGCGAAGGCCCTCAAGAC 791
    |||
QY 810 CACAGTCATCTCATCTGCGCTTCTTCCGCTGTGGCTGTGCTTACTACATTGGATTCAG 869
    |||
Db 792 CACAGTCATCTCATCTGCGCTTCTTCCGCTGTGGCTGTGCTTACTACATTGGATTCAG 851
    |||
QY 870 CATGCACTCTCATCTCTCTGGAATTCATCAAGCAAGGCTGACTTTGAGAACACTGT 929
    |||
Db 852 CATGCACTCTCTCATCTCTCTGGAATTCATCAAGCAAGGCTGACTTTGAGAACACTGT 911
    |||
QY 930 GCACAAGTGAATTTCCATCAGCAGGCGCTAGCTTCTTCTCAGCTGTTGTGTGAACCCAT 989
    |||
Db 912 GCACAAGTGAATTTCCATCAGCAGGCGCTAGCTTCTTCTCAGCTGTTGTGTGAACCCAT 971
    |||
QY 990 CCTCTATGCTTTCCTTGGAGCCAAATTTAAACCTCTGCCAGACGACACTCCTCTGT 1049
    |||
Db 972 CCTCTATGCTTTCCTTGGAGCCAAATTTAAACCTCTGCCAGACGACACTCCTCTGT 1031
    |||
QY 1050 GAGCAGAGGAGTCCAGGCTCAAGATCTCTCAGAAAGGAAAGCGAGTGGACATTCATCTGT 1109
    |||

```

```

Db 1032 GAGCAGAGGAGTCCAGCCTCAAGATCCTCTCCAAAGGAAGGAGGTGACATTCATCTGT 1091
QY 1110 TTCCTAGTCTGAGTCTTCAAGTTTTCACCTCCAGCTAACACAGATGTAAGAAGCTTTT 1169
    |||
Db 1092 TTCCTAGTCTGAGTCTTCAAGTTTTCACCTCCAGCTAACACAGATGTAAGAAGCTTTT 1151
    |||
QY 1170 TTTTATGCAATTAATACCTTTTATAGTACACATTTTCAGATTAATAAGAGCTGACC 1229
    |||
Db 1152 TTTTATGCAATTAATACCTTTTATAGTACACATTTTCAGATTAATAAGAGCTGACC 1211
    |||
QY 1230 AATATTGTACAGTTTATATGCTTGTGGATTTTGG-TCTTGTGTTCTTTAGTTT-T 1287
    |||
Db 1212 AATATTGTACAGTTTATATGCTTGTGGATTTTGGATTTTGTCTGTGTTCTTTAGTTTTCG 1271
    |||
QY 1288 TGAA-GTTTAAATGACTTATTTATTAATTTTGTGTTTTCATATTGATGTGTCTAG 1346
    |||
Db 1272 TGAAGGTTTAAATGACTTATTTATTAATTTTGTGTTTTCATATTGATGTGTCTAG 1331
    |||
QY 1347 GCAGAGACTGTGGCCAAAGTCTTAGTGTGCTGTATGCTGTGGTGTAGGACTGTGAAGAAAG 1406
    |||
Db 1332 GCAGAGACTGTGGCCAAAGTCTTAGTGTGCTGTATGCTGTGTGTGTGTGTGTGTGTGTGT 1391
    |||
QY 1407 GAACTGAACATTCAGAGCGGTGATGTAATCAGCTAAAGCTAGAAATGATCCCAAGCTGT 1466
    |||
Db 1392 GAACTGAACATTCAGAGCGGTGATGTAATCAGCTAAAGCTAGAAATGATCCCAAGCTGT 1451
    |||
QY 1467 TTAATGATATATATCTCTCATCTCCCGTGGAGACGTTTTTCTCTGTTCTTAAGACGTGAT 1526
    |||
Db 1452 TTAATGATATATATCTCTCATCTCCCGTGGAGACGTTTTTCTCTGTTCTTAAGACGTGAT 1511
    |||
QY 1527 TTGCTGTAGAAATGAGCAGCTTATTAACCAAGGCCCAAGGTGTATGAAGATCGTGGTTTTT 1586
    |||
Db 1512 TTGCTGTAGAAATGAGCAGCTTATTAACCAAGGCCCAAGGTGTATGAAGATCGTGGTTTTT 1571
    |||
QY 1587 CAGTTTTACAGAGTGGTGTATTTTCAGACCTAGACGTGATGATGATTTGATTAAGTTGTT 1646
    |||
Db 1572 CAGTTTTACAGAGTGGTGTATTTTCAGACCTAGACGTGATGATGATTTGATTAAGTTGTT 1631
    |||
QY 1647 AATTAAGTATACGTTAACTTAAAAAATAAAAA 1685
    |||
Db 1632 AATTAAGTATACGTTAACTTAAAAAATAAAAA 1670
    |||

```

RESULT 13
 HUMNYRECA
 LOCUS
 DEFINITION
 ACCESSION
 NID
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1225)
 Jazlin,E.E., Yoo,H., Blomqvist,A.G., Yee,F., Weng,G., Walker,M.W.,
 Salom,J., Lathamar,D. and Wahlestedt,C.
 A proposed bovine neuropeptide Y (NPY) receptor cDNA clone, or its
 human homologue, confers neither NPY binding sites nor NPY
 responsiveness on transfected cells
 Regul. Pept. 47 (3), 247-258 (1993)
 JOURNAL
 MEDLINE
 FEATURES
 SOURCE
 Location/Qualifiers
 1..1225
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 26..1084
 /gene="NPYR"
 26..1084
 /gene="NPYR"

```

/note="putative"
/codon_start=1
/product="neuropeptide Y receptor"
/protein_id="AAB1594.1"
/db_xref="PID:q189314"
/db_xref="GI:189314"
/translation="MEGISITVSDNTYEMSGDYDSMKPECFREANFNKIFLEPTI
YSIIFLTGIVNGLVILVMGYOKLRISMTDKYRLHLSVADLLFVITLPPWADVAMN
YFGNFLCAVHYIVTVNLYSVSLIAFISLDRYLAIVHATNSQPRKILAEKVYVGV
WIPALLITIPDFEAVNSEADRYICDRFYNIDMVVFOHIMVGLIPGIVILSC
YCIITSLKSHSGHOKRKALKTTVILIAFPACWLPYIGISIDSIILEIIKOGCEP
ENTVHKWISITELALAFHCCLNPIILYAFGLAKFKTSQAHALISVSGSSLLKSKGR
GGHSVSTSESSSFHSS"
BASE COUNT      288 a      324 c      259 g      354 t
ORIGIN
Query Match      63.8%; Score 1108; DB 9; Length 1225;
Best Local Similarity 100.0%; Pred.No.0;
Matches 1108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 56 CGCATCTGAGAACCCAGCGGTCTACATGAGGAGGATGATATATACCTTCAGATACTA 125
    |||
Db 1 CGCATCTGAGAACCCAGCGGTCTACATGAGGAGGATGATATATACCTTCAGATACTA 60
QY 126 CACCGAGGAATGGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGAGA 185
    |||
Db 51 CACCGAGGAATGGGCTCAGGGGACTATGACTCCATGAGGAACCCCTGTTCCGTGAGA 120
QY 186 AATGCTAATTTCAATTAATCTTCCGCCACCATCTACTCCATCTTTCTTAAGTG 245
    |||
Db 121 AATGCTAATTTCAATTAATCTTCCGCCACCATCTACTCCATCTTTCTTAAGTG 180
QY 246 CATTTGGGCAATGATTTGCTATCTGCTATGAGGTTTACAGAGAATGAGGAACAT 305
    |||
Db 181 CATTTGGGCAATGATTTGCTATCTGCTATGAGGTTTACAGAGAATGAGGAACAT 240
QY 306 GACGAGCAATGAGGCTGACCTGTGAGTGGCCGACCTCTCTTTGTCATCAGCCTTC 365
    |||
Db 241 GACGAGCAATGAGGCTGACCTGTGAGTGGCCGACCTCTCTTTGTCATCAGCCTTC 300
QY 366 CTCTGGGCAATGATTTGCTATCTGCTATGAGGTTTACAGAGAATGAGGAACAT 425
    |||
Db 301 CTCTGGGCAATGATTTGCTATCTGCTATGAGGTTTACAGAGAATGAGGAACAT 360
QY 426 CCATGTCATGACAGTCAACCTCTACAGAGAGTCCCTGATCTGCTTCATCAGTCT 485
    |||
Db 361 CCATGTCATGACAGTCAACCTCTACAGAGAGTCCCTGATCTGCTTCATCAGTCT 420
QY 486 GACCGCTACCTGGCCATCTGTCACGCCACCAACAGTCAAGGCCAAGAGCTGTTGGC 545
    |||
Db 421 GACCGCTACCTGGCCATCTGTCACGCCACCAACAGTCAAGGCCAAGAGCTGTTGGC 480
QY 546 TGAAGAAGTGTCTATGTTGGCGTCTGATCCCTGCTCTGCTGCTATTCGCCACTT 605
    |||
Db 481 TGAAGAAGTGTCTATGTTGGCGTCTGATCCCTGCTCTGCTGCTATTCGCCACTT 540
QY 606 CATCTTGGCAAGCTGACAGGAGAGATGAGATATATCTGTACCGCTTCATACCCCAA 665
    |||
Db 541 CATCTTGGCAAGCTGACAGGAGAGATGAGATATATCTGTACCGCTTCATACCCCAA 600
QY 666 TGACTTGTGGGTGGTGTGTTCCAGTTTCAGACATCATGTTGGCTTATCTGCTGG 725
    |||
Db 601 TGACTTGTGGGTGGTGTGTTCCAGTTTCAGACATCATGTTGGCTTATCTGCTGG 660
QY 726 TATTTGATCTCTGCTGCTATGATGATATATCTCCAAAGCTGTCAACATCCAAAGGCCA 785
    |||
Db 661 TATTTGATCTCTGCTGCTATGATGATATATCTCCAAAGCTGTCAACATCCAAAGGCCA 720
QY 786 CCGAAGCCGAAAGGCGCTCAAGACACAGATCATCTATCTGAGCTTTCTGGCTGTG 845
    |||
Db 721 CCGAAGCCGAAAGGCGCTCAAGACACAGATCATCTATCTGAGCTTTCTGGCTGTG 780

```

```

QY 846 GCTGCCCTTACTACATTTGGGATCAGATCGATCTCTTCACTCTCTGGAATCATCAAGA 905
    |||
Db 781 GCTGCCCTTACTACATTTGGGATCAGATCGATCTCTTCACTCTCTGGAATCATCAAGA 840
QY 906 AGGCTGTGATTTGGAACACGTGTGACAGAGTGGATTTCCATCACCAGGCGCTGCTTT 965
    |||
Db 841 AGGCTGTGATTTGGAACACGTGTGACAGAGTGGATTTCCATCACCAGGCGCTGCTTT 900
QY 966 CTTCACATGTTGCTGAAACCCCATCTCTATCTCTTCTGAGGCCAATTTAAACCTC 1025
    |||
Db 901 CTTCACATGTTGCTGAAACCCCATCTCTATCTCTTCTGAGGCCAATTTAAACCTC 960
QY 1026 TGCCAGCAGCAGCTACCTCTGTGAGCAGAGAGGTTCACAGCTCAAGATCTCTCAAAAG 1085
    |||
Db 961 TGCCAGCAGCAGCTACCTCTGTGAGCAGAGAGGTTCACAGCTCAAGATCTCTCAAAAG 1020
QY 1086 AAGCGAGGTGAGACATTCATCTGTTCCAGTACGTGCTCAAGTTTCACTCCAG 1145
    |||
Db 1021 AAGCGAGGTGAGACATTCATCTGTTCCAGTACGTGCTCAAGTTTCACTCCAG 1080
QY 1146 CTACACAGATGTAAAGACTTTT 1173
    |||
Db 1081 CTACACAGATGTAAAGACTTTT 1108
RESULT 14
AF025375
LOCUS AF025375 1059 bp mRNA PRI 28-OCT-1997
DEFINITION Homo sapiens chemokine receptor-4 (CXCR4) mRNA, complete cds.
ACCESSION AF025375
NID 92565335
VERSION AF025375.1 GI:2565335
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1059)
AUTHORS Xiao,L., Weiss,S., Qari,S., Rudolph,D., Hodge,T. and Lal,R.
TITLE Partial resistance to infection by syncytium-inducing primary HIV-1
in exposed uninfected individuals homozygous for CCR5 32bp deletion
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1059)
AUTHORS Qari,S.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-1997) Retrovirus Diseases Branch, Centers for
Diseases Control and Prevention, 1600 Clifton Road, Atlanta, GA
30333, USA
FEATURES
source location/Qualifiers
1..1059
/organism="Homo sapiens"
/isolate="Patient EUD32-1"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q21"
1..1059
/gene="CXCR4"
1..1059
/feature="CXCR4"
/note="fusin; G protein-coupled receptor; HUMSTR; LESTR"
/codon_start=1
/product="chemokine receptor-4"
/protein_id="AAB81970.1"
/db_xref="PID:92565336"
/db_xref="GI:2565336"
/translation="MEGISITVSDNTYEMSGDYDSMKPECFREANFNKIFLEPTI
YSIIFLTGIVNGLVILVMGYOKLRISMTDKYRLHLSVADLLFVITLPPWADVAMN
YFGNFLCAVHYIVTVNLYSVSLIAFISLDRYLAIVHATNSQPRKILAEKVYVGV
WIPALLITIPDFEAVNSEADRYICDRFYNIDMVVFOHIMVGLIPGIVILSC
YCIITSLKSHSGHOKRKALKTTVILIAFPACWLPYIGISIDSIILEIIKOGCEP
ENTVHKWISITELALAFHCCLNPIILYAFGLAKFKTSQAHALISVSGSSLLKSKGR
GGHSVSTSESSSFHSS"
BASE COUNT      244 a      301 c      232 g      282 t

```

ORIGIN

Query Match 61.0%; Score 1059; DB 11; Length 1059;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1059; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 91 ATGGAGGGGATCGATATATACACTCAGATACATACACCGAGAAATGGGCTCAGGGAC 150
    |||||||
Db 1 ATGGAGGGGATCGATATATACACTCAGATACATACACCGAGAAATGGGCTCAGGGAC 60

QY 151 TATGACTCAGTGAAGAACCCCTGTTCCGTGAAGAAAAATGCTAATTTCAATAAATCTTC 210
    |||||||
Db 51 TATGACTCAGTGAAGAACCCCTGTTCCGTGAAGAAAAATGCTAATTTCAATAAATCTTC 120

QY 211 CTGCCCCACCATCTACTCATTCTTCTTACTGCGATTTGTGGCAATGATTTGGTATC 270
    |||||||
Db 121 CTGCCCCACCATCTACTCATTCTTCTTACTGCGATTTGTGGCAATGATTTGGTATC 180

QY 271 CTGCTATGGGTTACAGAGAAACTGAGAAGCATGACGGACAAGTACAGGCTGACCTG 330
    |||||||
Db 181 CTGCTATGGGTTACAGAGAAACTGAGAAGCATGACGGACAAGTACAGGCTGACCTG 240

QY 331 TCATGGCCGACCTCCTCTTTGTATCAGCCTTCCCTTGGGCAATTGAATGCCGTGGCA 390
    |||||||
Db 241 TCATGGCCGACCTCCTCTTTGTATCAGCCTTCCCTTGGGCAATTGAATGCCGTGGCA 300

QY 391 AACTGTACTTTGGGAACCTCTATGCAAGGACGTCCATGCTATCTACAGTCAACCTC 450
    |||||||
Db 301 AACTGTACTTTGGGAACCTCTATGCAAGGACGTCCATGCTATCTACAGTCAACCTC 360

QY 451 TACAGCAGTGTCTCTATCTCTGACCTTATCATGCTGAGACCCGCTACCTGGCCATCTGCAC 510
    |||||||
Db 351 TACAGCAGTGTCTCTATCTCTGACCTTATCATGCTGAGACCCGCTACCTGGCCATCTGCAC 420

QY 511 GCCACCAAGAGTAGAGGCCAAGAAAGCTGTGGCTGAAAAAGTGGTCTATTTGGGCTC 570
    |||||||
Db 421 GCCACCAAGAGTAGAGGCCAAGAAAGCTGTGGCTGAAAAAGTGGTCTATTTGGGCTC 480

QY 571 TGGATCCCTGCCCTCTCTGCTGACTATTCGCGACTTATCTTCCCAAGCTCAGTGAAGCA 630
    |||||||
Db 481 TGGATCCCTGCCCTCTCTGCTGACTATTCGCGACTTATCTTCCCAAGCTCAGTGAAGCA 540

QY 631 GATGACAGATATATCTGTGACCCCTTCTACCCCAATGACTTGTGGGTGTGTGTTCCAG 690
    |||||||
Db 541 GATGACAGATATATCTGTGACCCCTTCTACCCCAATGACTTGTGGGTGTGTGTTCCAG 600

QY 691 TTTCACCAATATGTTGGCTTATCTGCGCTGATTTGCTCATCTCTGCTCTGCTATGC 750
    |||||||
Db 601 TTTCACCAATATGTTGGCTTATCTGCGCTGATTTGCTCATCTCTGCTCTGCTATGC 660

QY 751 ATTATCATCTCCAAGCTGTACACTCCAAAGGGGCCACAGAGGCAAGGCCCTCAAGACC 810
    |||||||
Db 661 ATTATCATCTCCAAGCTGTACACTCCAAAGGGGCCACAGAGGCAAGGCCCTCAAGACC 720

QY 811 ACAGTATCTCTATCTGCTGCTTCTTCCGCTGTTGGCTGCTTACTACATTTGGGATCAGC 870
    |||||||
Db 721 ACAGTATCTCTATCTGCTGCTTCTTCCGCTGTTGGCTGCTTACTACATTTGGGATCAGC 780

QY 871 ATGACTCTCTATCTGCTGCTGGAATCATCAAGAGGGGTGAGTTTGGAGACATGTTG 930
    |||||||
Db 781 ATGACTCTCTATCTGCTGCTGGAATCATCAAGAGGGGTGAGTTTGGAGACATGTTG 840

QY 931 CACAAGTGAATTCATCACCAGGCGCTAGCTTTCTTCCACTGTTGTGTGAAGCCCATC 990
    |||||||
Db 841 CACAAGTGAATTCATCACCAGGCGCTAGCTTTCTTCCACTGTTGTGTGAAGCCCATC 900

QY 991 CTCTATGCTTCTGCTGAGCAAAATTTAAAGCTGCGCAGACACCTCAGCTCTGTG 1050
    |||||||
Db 901 CTCTATGCTTCTGCTGAGCAAAATTTAAAGCTGCGCAGACACCTCAGCTCTGTG 960

QY 1051 ACCAGAGGGTCCAGCCTCAAGATCTCTCCAAAGAAAGGAGGTGAGACATTCATCTGT 1110
    |||||||
  
```

Db 961 AGCAGAGGTCACGCTCAAGATCCTCTCCAAAGAAAGCAGAGTGGACATTCATCTGTT 1020

QY 1111 TCCACTGAGTCTGAGTCTTCAAGTTTCACTCCAGCTAA 1149
 |||||||

Db 1021 TCCACTGAGTCTGAGTCTTCAAGTTTCACTCCAGCTAA 1059

RESULT 15

PTU89798 1087 bp mRNA PRI 23-FEB-1998

LOCUS Pan troglodytes CXK chemokine receptor 4 (CXCR4) mRNA, complete cds.

DEFINITION

ACCESSION U89798

NID 92281446

VERSION U89798.1 GI:2281446

KEYWORDS

SOURCE chimpanzee.

ORGANISM Pan troglodytes

REFERENCE 1 (bases 1 to 1087)
 Pretet,J.L., Zeribb,A.C., Girard,M., Guillet,J.G. and Butor,C.
 Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1

AUTHORS

JOURNAL AIDS Res. Hum. Retroviruses 13 (18), 1583-1587 (1997)

MEDLINE 98090115

REFERENCE 2 (bases 1 to 1087)
 Pretet,J.-L., Zeribb,A., Girard,M., Guillet,J.-G. and Butor,C.
 Direct Submission Submitted (17-FEB-1997) U445, INSERM, ICGM, 22 rue Mechain, Paris 75014, France

TITLE JOURNAL

FEATURES

source Location/Qualifiers

1..1087

/organism="Pan troglodytes"

/db_xref="taxon:9598"

/cell_type="PBL"

1..1087

/gene="CXCR4"

29..1087

/note="CXCR4"

/note="CXCR4: a seven-transmembrane domain receptor"

/codon_start=1

/product="CXK chemokine receptor 4"

/protein_id="AAC03718.1"

/db_xref="GI:2281447"

/db_xref="PID:92281447"

/translation="MEGISITYSDNTEMGSGDYDSMKPECFREBNAMNKIFLPTI
 VSILFLTGIVGIVLWVGXOKKIRSMTDKRLHLSVADLLFVILIPWAVDAVNM
 YGNTFLCAVHYITVNLKSSVLILAFISLDRLAIVAHNSORPKLLAEVYVYG
 WIPALDLTPDFTFANVSEADRVICDRYPNDLWVVFQFQIMVGLILPQIVILSC
 YCIILSKISHSGHOKRKALKTTVLILIAFFACWLPVYIGISIDSLILEITRKGEF
 ENTVHKMISITELALAFPHCCLENPLIYAFILGAFKFTSAOHALTSVSGSLIKLSGKR
 GHSVSSTSESSSFHSS"

BASE COUNT 250 a 311 c 239 g 287 t

ORIGIN

Query Match 56.7%; Score 985; DB 11; Length 1087;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1085; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 63 CACCGCATCTGGGAACACAGCGGTATACATGAGGGGATCAGTATATACACTTCAGATA 122
    |||||||
Db 1 CACCGCATCTGGGAACACAGCGGTATACATGAGGGGATCAGTATATACACTTCAGATA 60

QY 123 CTACACCGAAGAAATGGGCTCAGGGAGCTATGACTCATGAGAAGACCCGTTCGGTGA 182
    |||||||
Db 61 CTACACCGAAGAAATGGGCTCAGGGAGCTATGACTCATGAGAAGACCCGTTCGGTGA 120

QY 183 AGAAAAATGCTAATTTCAATAAATCTTCTGCCCCACCATCTACTCATCATATCTTTTAC 242
    |||||||
Db 121 AGAAAAATGCTAATTTCAATAAATCTTCTGCCCCACCATCTACTCATCATATCTTTTAC 180

QY 243 TGGCATTTGGGGCAATGGATTTGGTCTATCTGCTGATCTGGGTTACCAAGAAACTGAGAG 302
    |||||||
  
```

```

|||||
Db 181 TGGCATGTGGGCAATGATGGTGCATCTGGTGCATGGTTACAGAAAGCAAGTGAAG 240
QY 303 CATGACGGAAGAGACAGGCTGCACCTGTAGTGGCCGACCTCCTTTGTGATCAGCT 362
Db 241 CATGACGGAAGAGACAGGCTGCACCTGTAGTGGCCGACCTCCTTTGTGATCAGCT 300
QY 363 TCCCTTCTGGGCAAGTGTATCCCGTGGCAAACTGGTACTTTGGAACTTCTTATGCAAGG 422
Db 301 TCCCTTCTGGGCAAGTGTATCCCGTGGCAAACTGGTACTTTGGAACTTCTTATGCAAGG 360
QY 423 AGTCATGTATCTACACAGTCAACCTCTACAGAGTGTCTCATCTGGCCTTCATCAG 482
Db 361 AGTCATGTATCTACACAGTCAACCTCTACAGAGTGTCTCATCTGGCCTTCATCAG 420
QY 483 TCTGGACCGGTACTGGCCATTCGTCCACGCCACCAAGTCAGAGGCAAGAGTGT 542
Db 421 TCTGGACCGGTACTGGCCATTCGTCCACGCCACCAAGTCAGAGGCAAGAGTGT 480
QY 543 GGCCTGAAAAGGTGTCTATGTGGCGCTGGATCCCTGCCCTCGCTGACTATTCGCGA 602
Db 481 GGCCTGAAAAGGTGTCTATGTGGCGCTGGATCCCTGCCCTCGCTGACTATTCGCGA 540
QY 603 CTTCATCTTTGCCAAGCTCAGTGAAGCAGATGACAGATATATCTGTGACCGCTTACCC 662
Db 541 CTTCATCTTTGCCAAGCTCAGTGAAGCAGATGACAGATATATCTGTGACCGCTTACCC 600
QY 663 CAATGACTTGTGGGTGTGTGTTCAGTTTCAGACATATGTTGGCTTATCCCTGCC 722
Db 601 CAATGACTTGTGGGTGTGTGTTCAGTTTCAGACATATGTTGGCTTATCCCTGCC 660
QY 723 TGGTATGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 782
Db 661 TGGTATGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
QY 783 CCACCGAAGGCGAAGGCCCTCAAGACACAGTCAATCTCTCTCTCTCTCTCTCTCTCT 842
Db 721 CCACCGAAGGCGAAGGCCCTCAAGACACAGTCAATCTCTCTCTCTCTCTCTCTCTCT 780
QY 843 TTGGCTGCTTACTATGATGGGATGAGATGAGATGAGTCTCTCTCTCTCTCTCTCTCTCT 902
Db 781 TTGGCTGCTTACTATGATGGGATGAGATGAGATGAGTCTCTCTCTCTCTCTCTCTCTCT 840
QY 903 GCAAGGGTGTGATTTGAGAACACTGTGCACAAAGTGGATTTCCATCACCAGGCCCTAGC 962
Db 841 GCAAGGGTGTGATTTGAGAACACTGTGCACAAAGTGGATTTCCATCACCAGGCCCTAGC 900
QY 963 TTTCTTCCACTGTGTGTGTGAACCCCATCTCTATGCTTTCTTCTGGAGCCAAATTTAAAC 1022
Db 901 TTTCTTCCACTGTGTGTGTGAACCCCATCTCTATGCTTTCTTCTGGAGCCAAATTTAAAC 960
QY 1023 CTCTGCCAGCAGCAGTCACTCTCTGTGAGCAGAGGGTCCAGGCTCAAGTCTCTCCAA 1082
Db 961 CTCTGCCAGCAGCAGTCACTCTCTGTGAGCAGAGGGTCCAGGCTCAAGTCTCTCCAA 1020
QY 1083 AGGAAGCAGAGTGTGACATTCATCTTTTCCACTGAGTCTGAGTCTTCAAGTTTCACTC 1142
Db 1021 AGGAAGCAGAGTGTGACATTCATCTTTTCCACTGAGTCTGAGTCTTCAAGTTTCACTC 1080
QY 1143 CAGCTAA 1149
Db 1081 CAGCTAA 1087

RESULT 16
MMU93311 1059 bp mRNA PRI 14-MAY-1997
LOCUS Macaca mulatta alpha chemokine receptor (CXCR4) mRNA, complete cds.
DEFINITION
ACCESSION U93311
VERSION 91934670
KEYWORDS U93311.1 GI:1934670
SOURCE rhesus monkey.

```

```

ORGANISM Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;
Macaca.
REFERENCE 1 (bases 1 to 1059)
AUTHORS Edinger, A.L., Amedee, A., Miller, K., Doranz, B.J., Endres, M.,
Sharon, M., Samson, M., Lu, Z.-H., Clements, J.E., Murphy-Corb, M.,
Peifer, S.C., Parmentier, M., Broder, C.C. and Doms, R.W.
TITLE Differential utilization of CCR5 by macrophage and T cell tropic
simian immunodeficiency virus strains
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 94 (8), 4005-4010 (1997)
MEDLINE 97268687
REFERENCE 2 (bases 1 to 1059)
AUTHORS Miller, K.A., Hauer, D.A. and Clements, J.E.
TITLE Rhesus macaque CXCR4
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1059)
AUTHORS Miller, K.A., Hauer, D.A. and Clements, J.E.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1997) Comparative Medicine, Johns Hopkins
University School of Medicine, 720 Rutland Ave., Traylor G-60,
Baltimore, MD 21205, USA

FEATURES
source
1..1059
Location/Qualifiers
/organism="Macaca mulatta"
/db_xref="taxon:9544"
/cell_type="PHA-biased normal PBMC"
1..1059
/gene="CXCR4"
1..1059
/note="human CXCR4 ortholog; fusin; LESTR; HUMSTR; LCR1; G
protein-coupled receptor"
/codon_start=1
/product="alpha chemokine receptor"
/protein_id="AAB54116.1"
/db_xref="PID:91934671"
/db_xref="GI:1934671"
translation="MEGISITSONTEEMSGSDYSIKPECFRENAHFRILPTI
YSLFNGIVNGVIVLMGYOEKRSWTDKRIHLSTADLEFPIITLPRVADVAMN
KFGNFKAVAVITVNIYSVLLAFISLDRITAIHATVSRPKRLAKVYVGV
WIPALLTLPDIFRASVEADRYICOREPNDLVVYFQHLNMGVILGLIIVLSC
YCIILSKLSHGKORALKRTVYLILAFPCWLPYIGISIDSLILEIKGCEP
ENYVHKWISITELALFFHCNPLIYLAFLAKFKYTSQHALTIVSRSSILKSKR
GHSVSSESSSESNHSS"

BASE COUNT 241 a 301 c 236 g 281 t
ORIGIN

Query Match 19.6%; Score 341; DB 10; Length 1059;
Best Local Similarity 98.9%; Pred. No. 1e-173;
Matches 741; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 346 CTCTTTGTATACGCTTCCTCTGGGCAAGTGTGATGCGGTGGCAAACTGGTACTTTGGG 405
Db 346 CTCTTTGTATACGCTTCCTCTGGGCAAGTGTGATGCGGTGGCAAACTGGTACTTTGGG 315
QY 256 CTCCTTGTATACGCTTCCTCTGGGCAAGTGTGATGCGGTGGCAAACTGGTACTTTGGG 315
Db 256 CTCCTTGTATACGCTTCCTCTGGGCAAGTGTGATGCGGTGGCAAACTGGTACTTTGGG 315
QY 406 AACTTCCATGCAAGGAGTGCATGCTACACAGACCACTTACAGCAGTCTCTC 465
Db 406 AACTTCCATGCAAGGAGTGCATGCTACACAGACCACTTACAGCAGTCTCTC 375
QY 316 AACTTCCATGCAAGGAGTGCATGCTACACAGACCACTTACAGCAGTCTCTC 375
Db 316 AACTTCCATGCAAGGAGTGCATGCTACACAGACCACTTACAGCAGTCTCTC 375
QY 466 ATCTGAGCCTTCATGCTGTGAGCGCTACCTGAGCAGTGTCCAGCAGCAGCAGTGTAG 525
Db 466 ATCTGAGCCTTCATGCTGTGAGCGCTACCTGAGCAGTGTCCAGCAGCAGCAGTGTAG 435
QY 526 AGGCCAAGAGAGCTGTGGCTGAAAAAGTGTGATGTTGGCGTGTGATCCTGCGCTC 585
Db 526 AGGCCAAGAGAGCTGTGGCTGAAAAAGTGTGATGTTGGCGTGTGATCCTGCGCTC 495
QY 586 CTGCTGACTATCCGAGCTTCATCTTTCACAGCAGTGTGAGGAGAGATGAGATATATC 645
Db 586 CTGCTGACTATCCGAGCTTCATCTTTCACAGCAGTGTGAGGAGAGATGAGATATATC 555

```


ACCESSION D86579
 NID 91468948
 VERSION D86579.1 GI:1468948
 KEYWORDS fusin (LESTR).
 SOURCE Macaca fascicularis cDNA to mRNA.
 ORGANISM Macaca fascicularis
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
 REFERENCE 1 (bases 1 to 1078)
 AUTHORS Tatsumi, M.
 TITLE Direct Submission
 JOURNAL Submitted (18-JUL-1996) to the DDBJ/EMBL/GenBank databases. Masashi
 Tatsumi, National Institute of Health, Department of Veterinary
 Science, Toyama-23-1, Shinjuku-Ku, Tokyo 162, Japan
 (E-mail: tatsu@nih.go.jp, Tel: 81-3-5285-1111, Fax: 81-3-5285-1179)
 2 (sites)
 REFERENCE 2
 AUTHORS Monkey CD4 and fusin are not species barrier for HIV-1 replication
 TITLE Unpublished (1996)
 JOURNAL
 FEATURES
 source
 1..1078
 Location/Qualifiers
 /organism="Macaca fascicularis"
 /db_xref="taxon:9541"
 11..1069
 /codon_start=1
 /product="Fusin (LESTR)"
 /protein_id="BAAL3126.1"
 /db_xref="PID:d1013813"
 /db_xref="PID:g1468949"
 /db_xref="GI:1468948"
 /translation="MEGISIYTSNDYTEEMSGSDYSIKEPCRENAHNRILPPTI
 YSIFLIGVINGVILVMGYOKRLSMIDKYLHLSVDLFLVITLPPWADVAVNM
 YFENFLCAVAVIYVNLVSSVLILAFISDRILAVIATNSQRPRLAEKVVYGV
 WIPALLITIDPFIASVSEADRYICDRFPNDLVWVVOFOHINWGLIPGVILSC
 YCIILSKLSKHOKKALKITVILILAFACMLPYIIGISIDSLILEIKOGCEP
 ENTVHKWISITELAFHFHCCLNPLILVAFGLAKRTISAQHALTSVSGSSILKLSKGR
 GHSVSTSESSPFS"
 BASE COUNT 244 a 304 c 243 g 287 t
 ORIGIN

Query Match 19.3%; Score 336; DB 9; Length 1078;
 Best Local Similarity 98.9%; Pred. No. 5.2e-171;
 Matches 736; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

351 TGTTCATACGCTTCCTTCTGGGCGAGTTGATGCCGTGCAAACTGTACTTTGGAACTT 410
 |||||||
 271 TGTTCATACGCTTCCTTCTGGGCGAGTTGATGCCGTGCAAACTGTACTTTGGAACTT 330
 |||||||
 411 CCTATGCAAGGAGTCCATGTCATCTACAGTCAACCTCTACAGCAGTGTCTCATCT 470
 |||||||
 331 CCTGTGCAAGGAGTCCATGTCATCTACAGTCAACCTCTACAGCAGTGTCTCATCT 390
 |||||||
 471 GGCCTTCATGCTGTGAGCGCTACCTGSCCATGTCACAGCCACCAAGTACAGAGCC 530
 |||||||
 391 GGCCTTCATGCTGTGAGCGCTACCTGSCCATGTCACAGCCACCAAGTACAGAGCC 450
 |||||||
 531 AAGGAAGCTTTGGCTGAAAGGTGCTATGTTGGCGTGTGATCCCTGCTGCT 590
 |||||||
 451 AAGGAAGCTTTGGCTGAAAGGTGCTATGTTGGCGTGTGATCCCTGCTGCT 510
 |||||||
 591 GACTATTCGCGAGCTTCATCTTTCGCCAAGCTCATGTAGGCGAGATGACAGATATATCTGTA 650
 |||||||
 511 GACTATTCGCGAGCTTCATCTTTCGCCAAGCTCATGTAGGCGAGATGACAGATATATCTGTA 570
 |||||||
 651 CCCTTCCTACCCCAATGACTTGTGGGTGTGTGTCCAGTTTCAGACATCATGTTGG 710
 |||||||
 571 CCCTTCCTACCCCAATGACTTGTGGGTGTGTGTCCAGTTTCAGACATCATGTTGG 630
 |||||||
 711 CCTATTCCTGCTGCTGATTTGTCTGCTGCTGCTATGTCATATCATCTCCAGCTGC 770
 |||||||
 631 CCTATTCCTGCTGCTGATTTGTCTGCTGCTGCTATGTCATATCATCTCCAGCTGC 690
 |||||||

QY 771 AACTCCAAAGGCGCCACAGAAAGCGCAAGGCCCTTCAAGACCTACATCTCTGCGC 830
 |||||||
 DB 691 AACTCCAAAGGCGCCACAGAAAGCGCGCCCTCAAGACCGATCATCTCTGCGC 750
 |||||||
 QY 831 TTCTTCGCTGTGGCTGCTTACTACATTTGGATCAGATGATCCCTCATCTCTCT 890
 |||||||
 DB 751 TTCTTCGCTGTGGCTGCTTACTACATTTGGATCAGATGATCCCTCATCTCTCT 810
 |||||||
 QY 891 GGAATATATACGAAGGCTGTAGTTTGAGAACACTGTGCACAGTGTATTCATCAC 950
 |||||||
 DB 811 GGAATATATACGAAGGCTGTAGTTTGAGAACACTGTGCACAGTGTATTCATCAC 870
 |||||||
 QY 951 CGAGGCCCTGCTTTTTCACAGTGTGTGAGAACCCCATCTCTATCTCTCTGAGC 1010
 |||||||
 DB 871 CGAGGCCCTGCTTTTTCACAGTGTGTGAGAACCCCATCTCTATCTCTCTGAGC 930
 |||||||
 QY 1011 CAATTTAAACCTGTGCCAGCAGCAGCTCACTCTGTGACAGAGGCTCCAGCTCAA 1070
 |||||||
 DB 931 CAATTTAAACCTGTGCCAGCAGCAGCTCACTCTGTGACAGAGGCTCCAGCTCAA 990
 |||||||
 QY 1071 GATCCTCTCCAAAGGAAGCGAGG 1094
 |||||||
 DB 991 GATCCTCTCCAAAGGAAGCGAGG 1014
 |||||||

RESULT 19
 LOCUS AF019378 1126 bp mRNA PRI 02-JAN-1999
 DEFINITION Cercopithecus aethiops G-protein coupled receptor (CXCR4) mRNA,
 complete cds.
 ACCESSION AF019378
 NID AF019378
 VERSION AF019378.1 GI:4102991
 KEYWORDS
 SOURCE African green monkey.
 ORGANISM Chlorocebus aethiops
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;
 Chlorocebus.
 1 (bases 1 to 1126)
 REFERENCE Holtkamp, N., Baier, M. and Werner, A.
 AUTHORS CXCR4 from African green monkey
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 1126)
 AUTHORS Holtkamp, N., Baier, M. and Werner, A.
 TITLE Submitted (14-AUG-1997) 6/3, Paul-Ehrlich-Institut,
 JOURNAL Paul-Ehrlich-Str. 51-59, Langen 65225, Germany
 FEATURES
 source
 1..1126
 Location/Qualifiers
 /organism="Chlorocebus aethiops"
 /db_xref="taxon:9534"
 1..1126
 /gene="CXCR4"
 68..1126
 /gene="CXCR4"
 /codon_start=1
 /product="G-protein coupled receptor"
 /protein_id="AA01638.1"
 /db_xref="PID:g102992"
 /db_xref="GI:4102991"
 /translation="MEGISIYTSNDYTEEMSGSDYSIKEPCRENAHNRILPPTI
 YSIFLIGVINGVILVMGYOKRLSMIDKYLHLSVDLFLVITLPPWADVAVNM
 YFENFLCAVAVIYVNLVSSVLILAFISDRILAVIATNSQRPRLAEKVVYGV
 WIPALLITIDPFIASVSEADRYICDRFPNDLVWVVOFOHINWGLIPGVILSC
 YCIILSKLSKHOKKALKITVILILAFACMLPYIIGISIDSLILEIKOGCEP
 ENTVHKWISITELAFHFHCCLNPLILVAFGLAKRTISAQHALTSVSGSSILKLSKGR
 GHSVSTSESSPFS"
 BASE COUNT 254 a 318 c 260 g 294 t
 ORIGIN

Query Match 18.0%; Score 313; DB 11; Length 1126;

Best Local Similarity 98.8%; Pred. No. 1.5e-158;
Matches 813; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
OY 272 TGGTATGGGTACCAAGAAAGTGAAGCATGACGACAAAGTACAGGCTGCACCTGT 331
    |||||||
Db 249 TGGTATGGGTACCAAGAAAGTGAAGCATGACGACAAAGTACAGGCTGCACCTGT 308
OY 332 CAGTGGCCGACCTCTCTTTGTCATCAGCCTTCCCTTGGGGAGTGAATGCCGTGGCAA 391
    |||||||
Db 309 CAGTGGCTACCTCTCTTTGTCATCAGCCTTCCCTTGGGGAGTGAATGCCGTGGCAA 368
OY 392 ACTGGTACTTGGGAATCTCTATGCAAGGAGTCATGTCATCTACAGAGTCAACCTCT 451
    |||||||
Db 359 ACTGGTACTTGGGAATCTCTATGCAAGGAGTCATGTCATCTACAGAGTCAACCTCT 428
OY 452 ACAGAGTGTCTCATCTCTGGGCTTATGATGTCAGGCTGACCGCTGACCTGGCATCTGCAG 511
    |||||||
Db 429 ACAGAGTGTCTCATCTCTGGGCTTATGATGTCAGGCTGACCGCTGACCTGGCATCTGCAG 488
OY 512 CCACCAACAGTCAGAGGCGCAAGAGCTGTGGCTGAAAGTGATCTATGTTGGCTCT 571
    |||||||
Db 489 CCACCAACAGTCAGAGGCGCAAGAGCTGTGGCTGAAAGTGATCTATGTTGGCTCT 548
OY 572 GGATCCCTGCCCTCTCTGCTGACTATCCGACTTCATCTTTGCCAAGCTCAGTAGGCGAG 631
    |||||||
Db 549 GGATCCCTGCCCTCTCTGCTGACTATCCGACTTCATCTTTGCCAAGCTCAGTAGGCGAG 608
OY 632 ATGACAGATATCTGTGACGCTTCTACCCCAATGACTTGGGGTGTGTGTTCCACT 691
    |||||||
Db 609 ATGACAGATATCTGTGACGCTTCTACCCCAATGACTTGGGGTGTGTGTTCCACT 668
OY 692 TTCACACATCATGTTGGCTTTATCTGCTGCTGATTTGTCACTCTGCTGCTATGCA 751
    |||||||
Db 559 TTCACACATCATGTTGGCTTTATCTGCTGCTGATTTGTCACTCTGCTGCTATGCA 728
OY 752 TTATCATCTCCAAGCTGTACACTCCAAAGGCGCACCAAGCGCAAGGCCCTTCAGACCA 811
    |||||||
Db 729 TTATCATCTCCAAGCTGTACACTCCAAAGGCGCACCAAGCGCGCTTCAGACCA 788
OY 812 CAGTATCTCTCATCTCTGCTTCTTGGCTGTGGCTGCTTACTACTATGGGATGAGA 871
    |||||||
Db 789 CAGTATCTCTCATCTCTGCTTCTTGGCTGTGGCTGCTTACTACTATGGGATGAGA 848
OY 872 TCGACTCTCTCATCTCTCTGGAATCATCAAGCAAGGCTGTGAGTTTGAACACATGTC 931
    |||||||
Db 849 TCGACTCTCTCATCTCTCTGGAATCATCAAGCAAGGCTGTGAGTTTGAACACATGTC 908
OY 932 ACAAGTGGATTTCCATCAACGAGGCGCTAGCTTTCTTCACAGTGTGTGAACCCCATCC 991
    |||||||
Db 909 ACAAGTGGATTTCCATCAACGAGGCGCTAGCTTTCTTCACAGTGTGTGAACCCCATCC 968
OY 992 TCTATGCTTTCTTGGAGCAAAATTTAAAACCTGTGCCAGAGCAGTCACTCTGTGA 1051
    |||||||
Db 969 TCTATGCTTTCTTGGAGCAAAATTTAAAACCTGTGCCAGAGCAGTCACTCTGTGA 1028
OY 1052 GCAGAGGCTCAGAGCTCAAGATCTCTCTCCAAAGGAAAGGAGG 1094
    |||||||
Db 1029 GCAGAGGCTCAGAGCTCAAGATCTCTCTCCAAAGGAAAGGAGG 1071
```

RESULT 20
AF031089 1059 bp mRNA PRI 14-OCT-1998
LOCUS Papio hamadryas anubis CXC chemokine receptor 4 (CXC4) mRNA,
DEFINITION complete cds.
ACCESSION AF031089
NID 92625093
VERSION AF031089.1 GI:2625093
KEYWORDS olive baboon,
SOURCE Papio hamadryas anubis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae;

```
REFERENCE 1 (bases 1 to 1059)  
AUTHORS Benton, P.A., Lee, D.R., and Kennedy, R.C.  
TITLE Sequence comparisons of non-human primate HIV-1 coreceptor  
homologues  
JOURNAL Mol. Immunol. 35 (2), 95-101 (1998)  
MEDLINE 98346785  
REFERENCE 2 (bases 1 to 1059)  
AUTHORS Benton, P.A. and Kennedy, R.C.  
TITLE Direct Submission  
JOURNAL Submitted (22-Oct-1997) Microbiology and Immunology, University of  
Oklahoma Health Sciences Center, 800 Research Parkway, Rm 458,  
Oklahoma City, OK 73104, USA  
FEATURES  
source  
1..1059  
/organism="Papio hamadryas anubis"  
/sub_species="anubis"  
/db_xref="taxon:9555"  
/cell_type="peripheral blood mononuclear cells"  
1..1059  
/gene="CXC4"  
1..1059  
/gene="CXC4"  
/codon_start=1  
/product="CXC chemokine receptor 4"  
/protein_id="AAC63831.1"  
/db_xref="pid:92625094"  
/db_xref="gi:2625094"  
/translation="MEGISITSDNYTEEMSGSDYSIKEDCFRENAHFRIFLPTI  
YSILFLGIVNGVILVMGYOKRLRSMTDYLRLISVADLLFVTLFPMAVDVAMN  
YFNGKIAVHYIVTVNLSSVLLAIFSLDRYLAIYATNSQRPRLIAKRVYVG  
WIPALLIIPDFEIVASVEADRDYICDFSLDRLVWVVOFHVIGLILGILSC  
YCIISKLSHSKQKRAKLTVTYLLIAFACWLPYIGISIFILEIKOCCEP  
ENTYHKMSTIDALAFHRCNPLIYATLQKFKRTSAQHALTSRSGSLKILSKGR  
GHSVSTSESSSPHSS"  
BASE COUNT 237 a 301 c 238 g 283 t  
ORIGIN  
Query Match 18.0%; Score 313; DB 11; Length 1059;  
Best Local Similarity 98.8%; Pred. No. 1.5e-158;  
Matches 813; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
OY 272 TGGTATGGGTACCAAGAAAGTGAAGCATGACGACAAAGTACAGGCTGCACCTGT 331  
    |||||||  
Db 182 TGGTATGGGTACCAAGAAAGTGAAGCATGACGACAAAGTACAGGCTGCACCTGT 241  
OY 332 CAGTGGCCGACCTCTCTTTGTCATCAGCCTTCCCTTGGGGAGTGAATGCCGTGGCAA 391  
    |||||||  
Db 242 CAGTGGCTACCTCTCTTTGTCATCAGCCTTCCCTTGGGGAGTGAATGCCGTGGCAA 301  
OY 392 ACTGGTACTTGGGAATCTCTATGCAAGGAGTCCATGTCATCTACAGAGTCAACCTCT 451  
    |||||||  
Db 302 ACTGGTACTTGGGAATCTCTATGCAAGGAGTCCATGTCATCTACAGAGTCAACCTCT 361  
OY 452 ACAGAGTGTCTCATCTCTGGGCTTATGATGTCAGGCTGACCGCTGACCTGGCATCTGCAG 511  
    |||||||  
Db 359 ACAGAGTGTCTCATCTCTGGGCTTATGATGTCAGGCTGACCGCTGACCTGGCATCTGCAG 421  
OY 512 CCACCAACAGTCAGAGGCGCAAGAGCTGTGGCTGAAAGTGATCTATGTTGGCTCT 571  
    |||||||  
Db 429 CCACCAACAGTCAGAGGCGCAAGAGCTGTGGCTGAAAGTGATCTATGTTGGCTCT 481  
OY 572 GGATCCCTGCCCTCTCTGCTGACTATCCGACTTCATCTTTGCCAGGCTCAGTAGGCGAG 631  
    |||||||  
Db 482 GGATCCCTGCCCTCTCTGCTGACTATCCGACTTCATCTTTGCCAGGCTCAGTAGGCGAG 541  
OY 632 ATGACAGATATCTGTGACGCTTCTACCCCAATGACTTGGGGTGTGTGTTCCACT 691  
    |||||||  
Db 542 ATGACAGATATCTGTGACGCTTCTACCCCAATGACTTGGGGTGTGTGTTCCACT 601  
OY 692 TTCACACATCATGTTGGCTTTATCTGCTGCTGATGTCATCTGCTGCTATGCA 751  
    |||||||
```


Db 602 TTCAGACATCATGTTGGCCTTATCCTGCTGTAFCGTATCCTGCTGCTATGCA 661

Qy 752 TTATCATCTCCAAAGTGTACACTCCAAAGGCCCAAGAGGCCCTCAAGACCA 811

Db 652 TTATCATCTCCAAAGTGTACACTCCAAAGGCCCAAGAGGCCCTCAAGACCA 721

Qy 812 CATCATCTCCAAAGTGTACACTCCAAAGGCCCAAGAGGCCCTCAAGACCA 871

Db 722 CGGTCACTCTCCAAAGTGTACACTCCAAAGGCCCAAGAGGCCCTCAAGACCA 781

Qy 872 TCGACTCTCTCCAAAGTGTACACTCCAAAGGCCCAAGAGGCCCTCAAGACCA 931

Db 782 TCGACTCTCTCCAAAGTGTACACTCCAAAGGCCCAAGAGGCCCTCAAGACCA 841

Qy 932 ACAAGTGTATTCATCCAGAGGCCCTAGCTTCTCTCCAGAGGCCCTAGCTTCTCTGA 991

Db 842 ACAAGTGTATTCATCCAGAGGCCCTAGCTTCTCTCCAGAGGCCCTAGCTTCTCTGA 901

Qy 992 TGTATGCTCTCTCCAGAGGCCCTAGCTTCTCTCCAGAGGCCCTAGCTTCTCTGA 1051

Db 902 TGTATGCTCTCTCCAGAGGCCCTAGCTTCTCTCCAGAGGCCCTAGCTTCTCTGA 961

Qy 1052 GCAAGAGGCTCCAGAGGCCCTAGCTTCTCTCCAGAGGCCCTAGCTTCTCTGA 1094

Db 962 GCAAGAGGCTCCAGAGGCCCTAGCTTCTCTCCAGAGGCCCTAGCTTCTCTGA 1004

RESULT 21

AF051906 1059 bp mRNA PRI 03-JUL-1998

LOCUS Cercocebus torquatus atys PUG chemokine receptor CXCR4 mRNA, complete cds.

ACCESSION AF051906

VERSION AF051906.1 GI:3135303

KEYWORDS

SOURCE sooty mangabey.

ORGANISM Cercocebus torquatus atys

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Cercocebus.

REFERENCE 1 (bases 1 to 1059)

AUTHORS Chen, Z., Gettle, A., Ho, D.D., and Marx, P.A.

TITLE Primary SIVM isolates use the CCR5 coreceptor from sooty mangabeys naturally infected in west Africa: a comparison of coreceptor usage of Primary SIVM, HIV-2, and SIVmac

JOURNAL Virology 246 (1), 113-124 (1998)

MEDLINE 98321155

REFERENCE 2 (bases 1 to 1059)

AUTHORS Chen, Z.

TITLE Direct Submission

JOURNAL Submitted (03-MAR-1998) Aaron Diamond AIDS Research Center, 455 1st Ave. 7th Floor, NYC, NY 10016, USA

FEATURES

source

1. 1059

/organism="Cercocebus torquatus atys"

/db_xref="taxon:9531"

/cell_type="lymphocytes"

/note="Sooty mangabey PUG; from uninfected animal"

1. 1059

/codon_start=1

/product="chemokine receptor CXCR4"

/protein_id="AAC39834.1"

/db_xref="PDB:Q3135304"

/db_xref="GI:3135304"

/translation="MEGISITSDNYTEMSSGDYDIKEPCFEKNHNRIFLPTI YSIIFLTGIVNGLVILVMGYOKLRSMTRYRLHLSVADLLFVITTFPNAVAVANM YFNGFLCAVAVIYTVNLVYSSVLILATISDRYLAIVAHNSOKPRKLLAEKYVVG WIPALLTIPGEIFASVSEADDRICDFEPNDLVVVFQFQHIMVGLIPGIDILSC YCIISKLSHGKGRKALKTVILILAFACWLPYIGISIDISFILLEIKGCEP ENTYHKMISTEALAFPHCCNPLIYAFILAKFKTSQAHALTSVRSRSLKLSGKR GGHSSVSTESSESPHSS"

CDS

241 a 299 c 234 g 285 t

ORIGIN

Query Match 13.8%; Score 239; DB 11; Length 1059;

Best Local Similarity 100.0%; Pred. No. 1.6e-118;

Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 856 TACATTGGATCAGCATGCACTCTCTCATCTCTCTGGAATCATCAAGAGGTGTAG 915

Db 766 TACATTGGATCAGCATGCACTCTCTCATCTCTCTGGAATCATCAAGAGGTGTAG 825

Qy 916 TTGAGAACACTGTGACAAAGTGGATTTCCATCAGAGGCCCTACTTCTTCCACTG 975

Db 826 TTGAGAACACTGTGACAAAGTGGATTTCCATCAGAGGCCCTACTTCTTCCACTG 885

Qy 976 TGTCTGAACCCATCTCTATGCTTCTTGAGCCAAATTTAAACCTCTGCCAGCAC 1035

Db 886 TGTCTGAACCCATCTCTATGCTTCTTGAGCCAAATTTAAACCTCTGCCAGCAC 945

Qy 1036 GCACTCACTCTGTGAGCAGAGGCTCCAGCTCAAGATCTCTCCAAAGAGCGAGG 1094

Db 946 GCACTCACTCTGTGAGCAGAGGCTCCAGCTCAAGATCTCTCCAAAGAGCGAGG 1004

RESULT 22

MM073740 1059 bp mRNA PRI 21-MAR-1997

LOCUS Macaca mulatta CXCR4 mRNA, complete cds.

ACCESSION MM073740

VERSION MM073740.1 GI:1771982

KEYWORDS

SOURCE rhesus monkey.

ORGANISM Macaca mulatta

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.

REFERENCE 1 (bases 1 to 1059)

AUTHORS Chen, Z., Zhou, P., Ho, D.D., Landau, N.R., and Marx, P.A.

TITLE Genetically divergent strains of simian immunodeficiency virus use CCR5 as a coreceptor for entry

JOURNAL J. Virol. 71 (4), 2705-2714 (1997)

MEDLINE 97213934

REFERENCE 2 (bases 1 to 1059)

AUTHORS Chen, Z., and Marx, P.A.

TITLE Direct Submission

JOURNAL Submitted (08-OCT-1996) AIDS Animal Model, Aaron Diamond AIDS Research Center, 455 1st Ave., 7th Floor, New York, NY 10016, USA

FEATURES

source

1. 1059

/organism="Macaca mulatta"

/strain="Indian macaque"

/db_xref="taxon:9544"

/cell_type="PBMC"

1. 1059

/gene="CXCR4"

1. 1059

/gene="CXCR4"

/codon_start=1

/product="CXCR4"

/protein_id="AAC51159.1"

/db_xref="PDB:Q1899056"

/db_xref="GI:1899056"

/translation="MEGISITSDNYTEMSSGDYDIKEPCFEKNHNRIFLPTI YSIIFLTGIVNGLVILVMGYOKLRSMTRYRLHLSVADLLFVITTFPNAVAVANM YFNGFLCAVAVIYTVNLVYSSVLILATISDRYLAIVAHNSOKPRKLLAEKYVVG WIPALLTIPGEIFASVSEADDRICDFEPNDLVVVFQFQHIMVGLIPGIDILSC YCIISKLSHGKGRKALKTVILILAFACWLPYIGISIDISFILLEIKGCEP ENTYHKMISTEALAFPHCCNPLIYAFILAKFKTSQAHALTSVRSRSLKLSGKR GGHSSVSTESSESPHSS"

CDS

241 a 300 c 236 g 282 t

BASE COUNT

ORIGIN

Matches 77: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 TCGAAGCAGTCCATGTCATCTACACAGTCAACCTCTACAGAGTGTCTCTCTG6CC 474
|||||
Db 397 TCGAAGCAGTCCATGTCATCTACACAGTCAACCTCTACAGAGTGTCTCTCTG6CC 456
|||||

QY 475 TTCATCAGTCTGACCG 491
|||||

Db 457 TTCATCAGTCTGACCG 473
|||||

RESULT 25
FCU63558 1170 bp mRNA MAM 06-APR-1998
LOCUS Felis catus CXCR-4 homolog mRNA, complete cds.
DEFINITION
ACCESSION U63558
NID 93025443
VERSION 063558.1 GI:3025443
KEYWORDS
SOURCE
ORGANISM cat.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Carnivora; Fissipedia; Felidae; Felis.
REFERENCE 1 (bases 1 to 1170)
AUTHORS Willett,B.J., Picard,L., Hosie,M.J., Turner,J.D., Adema,K. and Clapham,P.R.
TITLE Shared usage of the chemokine receptor CXCR4 by the feline and human immunodeficiency viruses
J. Virol. 71 (9), 6407-6415 (1997)
MEDLINE 97404646
REFERENCE 2 (bases 1 to 1170)
AUTHORS Willett,B.J.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-1996) Veterinary Pathology, University of Glasgow, Bearsden Road, Glasgow G61 1QH, UK
REFERENCE 3 (bases 1 to 1170)
AUTHORS Willett,B.J.
TITLE Direct Submission
JOURNAL Submitted (06-APR-1998) Veterinary Pathology, University of Glasgow, Bearsden Road, Glasgow G61 1QH, UK
REMARK Sequence update by submitter
On Apr 6, 1998 this sequence version replaced gi:1731648.
FEATURES
Location/Qualifiers
source
1..1170
/organism="Felis catus"
/strain="domesticus"
/db_xref="taxon:9685"
/cell_type="T lymphoblasts"
84..1145
/function="receptor for chemokine SDF-1"
/note="LESTR"
/codon_start=1
/product="CXCR-4 homolog"
/protein_id="AAC48852.1"
/db_xref="PID:g1842247"
/db_xref="GI:1842247"
/translation="MDGFRIPSDNYTEDDLGSGDYDSKMEPCFRENAHFNRIPLPT
VYSIIIFLTGIVNGLVILVWGYSKRLSMIDKRLHLSVADLFLVTLPLPWAADAVAN
WYRGLCKRAVHYITVNIYSSVLIATFISIDRYLAIVHATNSGRPKLLAEKVYVG
VWIPALLITIPDFIFANVREADGRITCDRFPSSSWLVVFOFQIMGLIPGVILS
CYCIIISKLSHSGYOKRRALKTKTVIILIAFFACWLPYIGISIDSLILEITKQGE
FESTVHKWISITLALAFPHCCINPILYAFLGAKFKTSAOHALTSVSRGSSILKILSKG
RGHSSVSTESSESSFRSS"

BASE COUNT 268 a 320 c 269 g 313 t

ORIGIN

Query Match 4.4%; Score 77; DB 3; Length 1170;
Best Local Similarity 100.0%; Pred. No. 7.7e-31;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 TCGAAGCAGTCCATGTCATCTACACAGTCAACCTCTACAGAGTGTCTCTCTG6CC 474
|||||

Db 411 TCGAAGCAGTCCATGTCATCTACACAGTCAACCTCTACAGAGTGTCTCTCTG6CC 470

QY 475 TTCATCAGTCTGACCG 491
|||||

Db 471 TTCATCAGTCTGACCG 487
|||||

RESULT 26
FCU92795 1140 bp mRNA MAM 14-APR-1997
LOCUS Felis catus fusin (CXCR4) mRNA, complete cds.
DEFINITION
ACCESSION U92795
NID 91935044
VERSION 092795.1 GI:1935044
KEYWORDS
SOURCE
ORGANISM cat.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Carnivora; Fissipedia; Felidae; Felis.
REFERENCE 1 (bases 1 to 1140)
AUTHORS Ierner,D.L. and Elder,J.H.
TITLE Feline CXCR4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1140)
AUTHORS Ierner,D.L. and Elder,J.H.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-1997) Molecular Biology, TSRI, 10666 North Torrey
Pines Road, La Jolla, CA 92037, USA
FEATURES
Location/Qualifiers
source
1..1140
/organism="Felis catus"
/db_xref="taxon:9685"
/cell_line="MCH5-4"
/clone="15"
/note="lymphocytic cell line"
1..1062
/gene="CXCR4"
1..1062
/gene="CXCR4"
/note="LESTR"
/codon_start=1
/product="fusin"
/protein_id="AAB51765.1"
/db_xref="PID:g1935045"
/db_xref="GI:1935045"
/translation="MDGFRIPSDNYTEDDLGSGDYDSKMEPCFRENAHFNRIPLPT
VYSIIIFLTGIVNGLVILVWGYSKRLSMIDKRLHLSVADLFLVTLPLPWAADAVAN
WYRGLCKRAVHYITVNIYSSVLIATFISIDRYLAIVHATNSGRPKLLAEKVYVG
VWIPALLITIPDFIFANVREADGRITCDRFPSSSWLVVFOFQIMGLIPGVILS
CYCIIISKLSHSGYOKRRALKTKTVIILIAFFACWLPYIGISIDSLILEITKQGE
FESTVHKWISITLALAFPHCCINPILYAFLGAKFKTSAOHALTSVSRGSSILKILSKG
RGHSSVSTESSESSFRSS"

BASE COUNT 257 a 312 c 255 g 314 t 2 others

ORIGIN

Query Match 4.4%; Score 77; DB 3; Length 1140;
Best Local Similarity 100.0%; Pred. No. 7.7e-31;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 TCGAAGCAGTCCATGTCATCTACACAGTCAACCTCTACAGAGTGTCTCTCTG6CC 474
|||||

Db 328 TCGAAGCAGTCCATGTCATCTACACAGTCAACCTCTACAGAGTGTCTCTCTG6CC 387
|||||

QY 475 TTCATCAGTCTGACCG 491
|||||

Db 388 TTCATCAGTCTGACCG 404
|||||

RESULT 27
HUMNRY92 95 bp mRNA PRI 07-FEB-1999
LOCUS Human mRNA for neuropeptide Y3 receptor, 5'UTR (sequence from the
DEFINITION

ACCESSION	D28433	5'cap to the start codon).
NTD	9461211	
VERSION	D28433.1	GI:461211
KEYWORDS	neuropeptide y3 receptor.	
SOURCE	Homo sapiens (library: U937/pKAL) lymphoma cell-line U937 cDNA to mRNA, clone HP00400.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
AUTHORS	1 (bases 1 to 95)	
TITLE	Kato, S.	
JOURNAL	Direct Submission	
REFERENCE	Submitted (03-FEB-1994) to the DDBJ/EMBL/GenBank databases. Seishi Kato, Sagami Chemical Research Center, Genetic Engineering Section 4-4-1 Nishi-Ohnuma, Sagamihara, Kanagawa 229, Japan (E-mail: bdn00121@biotech.net.com, Tel:0427-42-4791(ex.415), Fax:0427-49-7631)	
AUTHORS	2 (bases 1 to 95)	
JOURNAL	Kato, S.	
REFERENCE	Unpublished (1994)	
AUTHORS	3 (sites)	
TITLE	Kato, S., Sekine, S., Oh, S.W., Kim, N.S., Umezawa, Y., Abe, N., Yokoyama-Kobayashi, M. and Aoki, T.	
JOURNAL	Construction of a human full-length cDNA bank	
MEDLINE	Gene 150 (2), 243-250 (1994)	
COMMENT	95121910	
COMMENT	Submitted (03-Feb-1994) to DDBJ by:	
	Seishi Kato	
	Genetic Engineering Section	
	Sagami Chemical Research Center	
	4-4-1 Nishi-Ohnuma	
	Sagamihara, Kanagawa 229	
	Japan	
	Phone: 0427-42-4791	
	Fax: 0427-49-7631.	
FEATURES	Location/Qualifiers	
Source	1..95	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/cell_line="U937"	
	/clone_lib="U937/pKAL"	
	/tissue_type="lymphoma"	
	1..89	
	90..>95	
	/codon_start=1	
	/product="neuropeptide y3 receptor"	
	/protein_id="BA05799.1"	
	/db_xref="PID:d1006346"	
	/db_xref="PID:g4433177"	
	/db_xref="GI:4433177"	
	/translation="ME"	
BASE COUNT	21 a 24 c 33 g 17 t	
ORIGIN		
Query Match	4.0%; Score 70; DB 9; Length 95;	
Best Local Similarity	100.0%; Pident. No. 5,2e-27;	
Matches	70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	27 GCAAGTAGCCCGGAGGCGCTGAGTGCCTCCAGTAGCCACCGCATCTGGAGACCAACCGGT 86	
Db	26 GCAAGTAGCCCGGAGGCGCTGAGTGCCTCCAGTAGCCACCGCATCTGGAGACCAACCGGT 85	
QY	87 TACCATGGAG 96	
Db	86 TACCATGGAG 95	
RESULT	28	
LOCUS	BOVNYR 1280 bp mRNA	
DEFINITION	Cow neuropeptide y receptor mRNA sequence.	
	12-FEB-1992	

```

ACCESSION  M86739
NID         GI:63428
VERSION     M86739.1  GI:163428
KEYWORDS    neuropeptide y receptor.
SOURCE      Bos taurus Brain locus coeruleus cDNA to mRNA.
ORGANISM    Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
REFERENCE   1 (bases 1 to 1280)
AUTHORS    Bovinae, Bos.
            Rimland, J., Xin, W.W., Sweetnam, P., Saijoh, K., Nestler, E.J. and
            Duman, R.S.
TITLE       Sequence and expression of a neuropeptide y receptor cDNA
JOURNAL     Mol. Pharmacol. 40, 869-875 (1991)
MEDLINE     92100053
FEATURES
    source
        1..1280
            /organism="Bos taurus"
            /db_xref="taxon:9913"
            /tissue_type="Brain locus coeruleus"
BASE COUNT  284 a 333 c 290 g 373 t
ORIGIN
Query Match 3.6%; Score 62; DB 3; Length 1280;
Best Local Similarity 100.0%; Pred. No. 1e-22;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 415 TGCAGGCGAGTCATGTCTATCTACACAGTCACACCTCTCTACAGCAGTGTCTCTACTCTGCGC 474
      |||||||
Db 388 TGCAGGCGAGTCATGTCTATCTACACAGTCACACCTCTCTACAGCAGTGTCTCTACTCTGCGC 447
      |||
QY 475 TT 476
      ||
Db 448 TT 449
RESULT 29
MMU059760 LOCUS 1223 bp mRNA ROD 10-SEP-1996
DEFINITION Mus musculus fusin homolog mRNA, complete cds.
ACCESSION U59760
NID g1527134
VERSION 059760.1 GI:1527134
KEYWORDS
SOURCE
ORGANISM
    house mouse.
    Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1223)
AUTHORS    Heesen, M., Berman, M.A., Gerard, C. and Dorf, M.E.
TITLE       Cloning of the mouse homologue of the human HIV co-factor gene,
            fusin
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 1223)
AUTHORS    Heesen, M., Berman, M.A., Gerard, C. and Dorf, M.E.
TITLE       Direct Submission
JOURNAL     Submitted (02-JUN-1996) Pathology, Harvard Medical School, 200
            Longwood Avenue, Boston, MA 02115, USA
FEATURES
    source
        1..1223
            /organism="Mus musculus"
            /strain="129/Sv"
            /db_xref="taxon:10090"
            /note="seven transmembrane spanning receptor;
            chemokine-receptor like"
            /codon_start=1
            /product="fusin homolog"
            /protein_id="AA077725.1"
            /db_xref="PID:g1527135"
            /db_xref="GI:1527135"
            /translation="MEPISVSIYTSNDYSEEVGSGDYDSNKEPCFRDENVHFNRFLIP

```

11VF1FLTGIVNGVLIVLMGYOKLRKSTDKYRLHLVSADLLFYTLPLFNAVDAMA
DMYFGKFLCKAVHIIYTVNLVSSVLIILAFISLDRIYAIYHATNSQRPRLAEKAVY
GWWIPALLLITIPDIPEIFADVSQDISOGBRIDRLPYPSLWVVFQFQIHWGLLP
GIYILSCYCIILIKSHSGHOKRKALKTIVILIAFFACMLPYVIGISIDSPILIGV
IKOGCPESIVHKWISTITELAFHCCLPILYAFIAGAFKSSAQAHLNMSRSGSLK
TISKGRGHSVSSTESSSFHSS"
BASE COUNT 285 a 342 c 268 g 328 t
ORIGIN

Query Match 3.2%: Score 56; DB 12; Length 1223;
Best Local Similarity 100.0%; Pred. No. 1.8e-19;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 754 ATCATCTCCAGCTGTCCACACTCCAGGCGCCACAGAGCGCCCTCAAGAC 809
|||||
DB 714 ATCATCTCCAGCTGTCCACACTCCAGGCGCCACAGAGCGCCCTCAAGAC 769

RESULT 30
LOCUS MMU65580 3366 bp DNA ROD 14-DEC-1996
DEFINITION Mus musculus fusin (CXCR-4) gene, complete cds.
ACCESSION U65580
NID 91731650
VERSION U65580.1 GI:1731650
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Heesen, M., Berman, M.A., Benson, J.D., Gerard, C. and Dorf, M.E.
TITLE Cloning of the mouse fusin gene, homologue to a human HIV-1
co-factor.
JOURNAL J Immunol. 157 (12), 5455-5460 (1996)
MEDLINE 9711334
REFERENCE
AUTHORS Heesen, M., Berman, M.A., Gerard, C. and Dorf, M.E.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-1996) Pathology, Harvard Medical School, 200
Longwood Avenue, Boston, MA 02115, USA

FEATURES
source 1..3366
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
1..21
/gene="CXCR-4"
/number=1
1..15
/gene="CXCR-4"
/note="alternate exon 1"
1..3366
/gene="CXCR-4"
join(1..21,2308..3366)
/gene="CXCR-4"
/note="SDF-1 chemokine receptor"
/codon_start=1
/product="Fusin"
/protein_id="AAC52953.1"
/db_xref="GI:3645915"
/translation="MEPIVSIVTSDNYSEVSGSDYDSNKEPCFDENENRIFLPL
TIFIFLIVNGVILVLMGYOKLRKSTDKYRLHLVSADLLFYTLPLFNAVDAMA
DMYFGKFLCKAVHIIYTVNLVSSVLIILAFISLDRIYAIYHATNSQRPRLAEKAVY
GWWIPALLLITIPDIPEIFADVSQDISOGBRIDRLPYPSLWVVFQFQIHWGLLP
GIYILSCYCIILIKSHSGHOKRKALKTIVILIAFFACMLPYVIGISIDSPILIGV
IKOGCPESIVHKWISTITELAFHCCLPILYAFIAGAFKSSAQAHLNMSRSGSLK
TISKGRGHSVSSTESSSFHSS"
16..>2307
/gene="CXCR-4"
/note="alternate intron 1"

Intron

intron 22..2307
/gene="CXCR-4"
/number=1
2308..3366
/gene="CXCR-4"
/number=2
BASE COUNT 700 a 874 c 849 g 943 t
ORIGIN

Query Match 3.2%: Score 56; DB 12; Length 3366;
Best Local Similarity 100.0%; Pred. No. 1.7e-19;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 754 ATCATCTCCAGCTGTCCACACTCCAGGCGCCACAGAGCGCCCTCAAGAC 809
|||||
DB 2971 ATCATCTCCAGCTGTCCACACTCCAGGCGCCACAGAGCGCCCTCAAGAC 3026

RESULT 31
LOCUS RNU54791 1362 bp mRNA ROD 05-JUN-1996
DEFINITION Rattus norvegicus chemokine receptor LCR1 mRNA, complete cds.
ACCESSION U54791
NID 91354504
VERSION U54791.1 GI:1354504
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
AUTHORS Wong, M., Xin, W.W. and Duman, R.S.
TITLE Rat LCR1: Cloning and cellular distribution of a putative chemokine
receptor in brain
JOURNAL Molecular Psychiatry (1996) In press
MEDLINE 2 (bases 1 to 1362)
REFERENCE
AUTHORS Wong, M.
TITLE Direct Submission
JOURNAL Submitted (11-APR-1996) Ma-Li Wong, Clinical Neuroendocrinology
Branch, National Institutes of Health, NIMH, Building 10 Room
3S231, 10 Center DR MSC 1284, Bethesda, MD 20892-1284, USA

FEATURES
source 1..1362
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="RLCR1"
44..1042
/codon_start=1
/product="chemokine receptor LCR1"
/protein_id="AAB01981.1"
/db_xref="PID:91354505"
/db_xref="GI:1354505"
/translation="WEITSDNYSEVSGSDYDSNKEPCFDENENRIFLPLTIFL
IFLTVGIVNGVILVLMGYOKLRKSTDKYRLHLVSADLLFYTLPLFNAVDAMA
DMYFGKFLCKAVHIIYTVNLVSSVLIILAFISLDRIYAIYHATNSQRPRLAEKAVY
GWWIPALLLITIPDIPEIFADVSQDISOGBRIDRLPYPSLWVVFQFQIHWGLLP
GIYILSCYCIILIKSHSGHOKRKALKTIVILIAFFACMLPYVIGISIDSPILIGV
IKOGCPESIVHKWISTITELAFHCCLPILYAFIAGAFKSSAQAHLNMSRSGSLK
TISKGRGHSVSSTESSSFHSS"
LNPITATLGAFTKTSQAHLNMSRSGSLKTLISKGRGHSVSSTESSSFHSS"

CDS

BASE COUNT 316 a 346 c 320 g 380 t
ORIGIN

Query Match 3.2%: Score 56; DB 12; Length 1362;
Best Local Similarity 100.0%; Pred. No. 1.8e-19;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 754 ATCATCTCCAGCTGTCCACACTCCAGGCGCCACAGAGCGCCCTCAAGAC 809
|||||
DB 647 ATCATCTCCAGCTGTCCACACTCCAGGCGCCACAGAGCGCCCTCAAGAC 702

```

RESULT 32
RN090610 1050 bp mRNA ROD 26-MAR-1997
LOCUS Rattus norvegicus CXc chemokine receptor (CXCR4) mRNA, complete
DEFINITION
cds.
ACCESSION U90610
NID 91906612
VERSION U90610.1 GI:1906612
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 1050)
AUTHORS Harrison, J.K. and Salafianca, M.N.
TITLE Molecular cloning of rat CXCR4
JOURNAL Unpublished
2 (bases 1 to 1050)
REFERENCE Harrison, J.K. and Salafianca, M.N.
AUTHORS
TITLE
JOURNAL
DIRECT SUBMISSION
Submitted (25-FEB-1997) Pharmacology and Therapeutics, University
of Florida, 1600 SW Archer Rd., Gainesville, FL 32610-0267, USA
FEATURES
Source
Location/Qualifiers
1..1050
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/tissue_type="spleen"
1..1050
/gene="CXCR4"
1..1050
/gene="CXCR4"
/feature="G-protein coupled receptor"
/codon_start=1
/product="CXC chemokine receptor"
/protein_id="AAB50408.1"
/db_xref="PID:91906613"
/db_xref="GI:1906613"
/translacion="MEITYSDNYSSEVGGSDYDKNKEPCFDEENENRIFLPIITFI
ILFLIGNGCIVILVMGYOKLRSMTOGRHLISVADILFPIITLPEFAVDMAAMWTFI
KFLCKAVHIITVNLVSSVLLAFISLDRIATVHATNSOSARLLEKAYVGMWIP
ALLLTIPDIIPADVSQDGRICDRLVPSLMMVVFQHLVGLILPGIVILSCYCI
IISKLSKSHQKRAKLTIVILIAFFACWLPYVGISIDSFILLEVIKGCEFSY
VHKWISITELAFHCCINPLIYAFILGAKFKSSQHALNSMRSSSLKILSKRGKH
SSVSTESSESSFS"
BASE COUNT 231 a 313 c 247 g 259 t
ORIGIN
Query Match 3.2%; Score 56; DB 12; Length 1050;
Best Local Similarity 100.0%; Pred. No. 1.8e-19;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 754 ATCATCTCAAGCTGTCACTCCAGGCGCAGAAAGCGCAGCCCTCAAGAC 809
Db 655 ATCATCTCAAGCTGTCACTCCAGGCGCAGAAAGCGCAGCCCTCAAGAC 710
RESULT 33
LOCUS E07381 3581 bp RNA PAT 29-SEP-1997
DEFINITION cDNA encoding mouse OSF-4.
ACCESSION E07381
NID 92175520
VERSION E07381.1 GI:2175520
KEYWORDS JP 1994122700-A/1.
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3581)
AUTHORS Takeshita, A., Okazaki, M., Kawai, S., Tsujimura, A. and Aman, E.
TITLE BONE-RELATED CADHERIN-LIKE PROTEIN AND ITS PRODUCTION

```

```

JOURNAL Patent: JP 1994122700-A 1 06-MAY-1994;
COMMENT HOECHST JAPAN LTD
OS Mus musculus (mouse)
PN JP 1994122700-A/1
PD 06-MAY-1994
PR 13-JUL-1993 JP 1993172883
PF 28-AUG-1992 JP 92P 230028
PI TAKESHITA ATSUSHI, OKAZAKI MAKOTO, KAWAI SHINJI, PI
TSUJIMURA ATSUSHI,
PI AMAN EGON
PC C07K13/00,A61K37/02,A61K39/395,A61K49/00,
PC C07K15/14,
PC C12N15/12,C12P21/02,C12P21/08;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FH source 1..3581
FT /organism="Mus musculus"
FT /strain="MC3T3-E1"
FT /clone="pKOT1164"
FT 5'UTR 1..283
FT CDS 284..2674
FT /product="Mouse OSF-4"
FT 2675..3581.
FT 3'UTR
Location/Qualifiers
1..3581
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 1121 a 773 c 801 g 886 t
ORIGIN
Query Match 3.1%; Score 53; DB 5; Length 3581;
Best Local Similarity 100.0%; Pred. No. 7.2e-18;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1669 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGC 1721
Db 3523 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGC 3575
RESULT 34
LOCUS AF002985 995 bp mRNA PRI 04-MAR-1999
DEFINITION Homo sapiens putative alpha chemokine (H174) mRNA, complete cds.
ACCESSION AF002985
NID 92580585
VERSION AF002985.1 GI:2580585
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 995)
AUTHORS Jacobs, K.A., Collins-Racie, L.A., Colbert, M., Duckett, M.,
Golden-Fleet, M., Kelleher, K., Kitz, R., Lavallee, E.R., Merberg, D.,
Spaulding, V., Stover, J., Williamson, M.J. and McCoy, J.M.
A genetic selection for isolating cDNAs encoding secreted proteins
Gene 198 (1-2), 289-296 (1997)
MEDLINE 98036061
REFERENCE 2 (bases 1 to 995)
AUTHORS Luo, Y., Kim, R., Gabuzda, D., Mi, S., Collins-Racie, L.A., Lu, Z.,
Jacobs, K.A. and Dorf, M.E.
The CXc-chemokine, H174: expression in the central nervous system
J. Neurovirol. 4 (6), 575-585 (1998)
TITLE JOURNAL
MEDLINE 99163782
REFERENCE 3 (bases 1 to 995)
AUTHORS Jacobs, K.A., Collins-Racie, L.A., Colbert, M., Duckett, M.,
Golden-Fleet, M., Kelleher, K., Kitz, R., Lavallee, E.R., Merberg, D.,
Spaulding, V., Stover, J., Williamson, M.J. and McCoy, J.M.

```

TITLE Direct Submission
JOURNAL Submitted (07-MAY-1997) Genetics Institute, 87 Cambridge Park Drive, Cambridge, MA 02140, USA
FEATURES
SOURCE location/Qualifiers
1..995
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="PMA and PMA activated human peripheral blood mononuclear cells"
1..995
/gene="H174"
88..372
/gene="H174"
/codon_start=1
/product="putative alpha chemokine"
/protein_id="AAC51845.1"
/db_xref="PID:9280586"
/db_xref="GI:2580586"
/translation="MSYKGMALALAVILCATVVGFPMEKRGRCICIGYKAVKAVAD
IEKASIMPSNMCDEIVITLTKENKGGRCINPKSKQARLLIKKVERKNF"
BASE COUNT 382 a 170 c 194 g 249 t
ORIGIN

Query Match 3.1%; Score 53; DB 42; Length 995;
Best Local Similarity 100.0%; Pred. No. 7.6e-18;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1659 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCC 1721
Db 943 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCC 995

RESULT 35
OAU38942 578 bp mRNA MAM 14-NOV-1995
LOCUS OAU38942
DEFINITION Ovis aries neuropeptide Y receptor type 3 mRNA, partial cds.
ACCESSION U38942
NID g1061411
VERSION U38942.1 GI:1061411
KEYWORDS
SOURCE sheep.
ORGANISM Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;
Caprinae; Ovis.
1 (bases 1 to 578)
Dyer,C.J., Matteri,R.L. and Keisler,D.H.
Development of an ovine Y3 cDNA and expression of the Y3 receptor
mRNA in the ovine hypothalamus and pituitary
Abstr. - Soc. Neurosci. 21, 1890-1890 (1995)
2 (bases 1 to 578)
Dyer,C.J., Matteri,R.L. and Keisler,D.H.
Direct Submission
Submitted (19-OCT-1995) Cheryl J. Dyer, Animal Science, University
of Missouri, 160 ASHC, Columbia, MO 65211, USA
Location/Qualifiers
1..578
/organism="Ovis aries"
/db_xref="taxon:9940"
/tissue_type="hypothalamus"
1..578
/note="synonyms: LCRL, LESTR, Y3 receptor; putative
neuropeptide Y receptor"
/codon_start=3
/product="neuropeptide Y receptor type 3"
/protein_id="AA81347.1"
/db_xref="PID:g1061412"
/db_xref="GI:1061412"
/translation="YTEDDLSGDYDKMKPCRPENAHENRIFLPTVYSITFLGIV
GNGLVLMGVOKRLKRSMDKRYRLHLSADLLFLFLPFAVADAVANMYRGOFLCKAV
HVTIVLVSSVLAFISDRYLAIVHAITSQPRKLLAEKVIVYGVWDPVLLTTP
DLIRADIKADERICDRFIPSDLVLFQFO"

BASE COUNT 121 a 157 c 146 g 154 t
ORIGIN

Query Match 2.9%; Score 50; DB 3; Length 578;
Best Local Similarity 100.0%; Pred. No. 3.7e-16;
Matches 50; Conservative 0; Mismatches 0; Indels 0;

OY 442 CTCACCTCTACAGAGTGCTCTATCTGGCTTCATCATGTGACCG 491
Db 324 GTCAACCTCTACAGAGTGCTCTATCTGGCTTCATCATGTGACCG 373

RESULT 36
I88893/c 188893 60 bp DNA PAT 17-JUL-1998
LOCUS I88893
DEFINITION Sequence 11 from patent US 5719125.
ACCESSION I88893
NID g3408833
VERSION I88893.1 GI:3408833
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 60)
Suzuki,F., Hiraki,Y., Takahashi,K., Suzuki,J., Kondo,J., Kohara,A.,
Mori,A. and Yamada,E.
Human Chondromodulin-1 protein
Patent: US 5719125-A 11 17-FEB-1998;
Location/Qualifiers
1..60
/organism="unknown"

BASE COUNT 2 a 8 c 8 g 42 t
ORIGIN

Query Match 2.9%; Score 50; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.5e-16;
Matches 50; Conservative 0; Mismatches 0; Indels 0;

OY 1674 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCA 1723
Db 60 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCA 11

RESULT 37
D87747 1877 bp mRNA ROD 07-FEB-1999
LOCUS D87747
DEFINITION Mus musculus mRNA for murine CXCR-4, complete cds.
ACCESSION D87747
NID g1772445
VERSION D87747.1 GI:1772445
KEYWORDS
SOURCE murine CXCR-4.
ORGANISM Mus musculus
Mus musculus bone marrow pre-B cell cell_line:DW34 cDNA to mRNA.
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1877)
Nagasawa,T.
Direct Submission
Submitted (05-SEP-1996) to the DDBJ/EMBL/Genbank databases. Takashi
Nagasawa, Research Institute, Osaka Medical Center, Department of
Immunology; Murodochi 840, Izumi, Osaka 590-02, Japan
(E-mail:immunol@osk.threewebnet.or.jp, Tel:0725-561220,
Fax:0725-57-3021)
2 (bases 1 to 1877)
Nagasawa,T.
Nagasawa,T.
Molecular cloning of murine PBSF/SDF-1 receptor
Unpublished (1996)
3 (sites)
Nagasawa,T., Nakajima,T., Tachibana,K., Itazawa,H., Bleul,C.C.,
Yoshie,O., Matsushima,K., Yoshida,N., Springer,T.A. and
Kishimoto,T.

JOURNAL	Molecular cloning and characterization of a murine pre-B-cell growth-stimulating factor/stromal cell-derived factor 1 receptor, a murine homolog of the human immunodeficiency virus 1 entry coreceptor fusin
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 93 (25), 14726-14729 (1996)
FEATURES	97121456
source	Location/Qualifiers 1..1877 /organism="Mus musculus" /db_xref="taxon:10090" /cell_line="DW34" /cell_type="pre-B cell" /tissue_type="bone marrow" 120..1199 /gene="murine CXCR-4" 120..1199 /gene="murine CXCR-4" /function="PBSF/SDF-1 receptor" /codon_start=1 /evidence="experimental" /product="murine CXCR-4" /protein_id="PAA13451.1" /db_xref="PID:d1014141" /db_xref="PID:g1772446" /db_xref="GI:1772446" /translation="MEPIVSIVITSDNYSEEVGSGDYDSNKEPCFRDENVHFNRIPLFLTIVYIFLFTGVGGVLILVMGYOKRSMQDKRYRLHSVDLLFVTLFPWADAVAA DMVGRFKCAKHAIITYVNLSYSVLIAFISLDRYLAIVHATNSORPKLIAEAYVA GVMIPALLITPTDIPADVSGDISOGDRVICRILVPDSIMVMVPOFOHIMGLIIP GVILSVCLITIKSLSKSGHKRKAKTYILLILAFAACGLPYVGISIDSFTLLIV IKOCDFESIYHKRISTITEALAEFHCCNPILVLFCLKRFSNAOHALNSRSSSLK ILSGCKRGSHSVSTESSESSFHS"
CDS	BASE COUNT 469 a 450 c 407 g 551 t ORIGIN
gene	
Query Match	2.9% Score 50; DB 12; Length 1877; Best Local Similarity 100.0%; Pctd. No. 3.1e-16;
LOCUS	Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1828 TAAACTTA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1877
RESULT 38	
LOCUS	S81785 495 bp mRNA INV 21-JUL-1996
DEFINITION	Myr p II-major allergen [Myrmecia pilosula=Australian jumper ants, abdomens, mRNA partial, 495 nt].
ACCESSION	S81785
NID	G1438760
VERSION	S81785.1 GI:1438760
KEYWORDS	.
SOURCE	Australian jumper ant abdomens.
ORGANISM	Myrmecia pilosula
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pserygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Formicoidea; Formicidae; Myrmecia.
AUTHORS	I (bases 1 to 495) Street,M.D., Donovan,G.R. and Baldo,B.A.
TITLE	Molecular cloning and characterization of the major allergen Myr p II from the venom of the jumper ant Myrmecia pilosula: Myr p I and Myr p II share a common protein leader sequence
JOURNAL	Biochim. Biophys. Acta 1305 (1-2), 87-97 (1996)
MEDLINE	GenBank staff at the National Library of Medicine created this entry [NCBI gibseq 176826] from the original journal article.
REMARK	This sequence comes from Fig. 5. Location/Qualifiers 1..495 /organism="Myrmecia pilosula" /db_xref="taxon:13618"
FEATURES	
Source	

gene		44.. .271	/gene="Myr p II"
CDS		44.. .271	/note="major allergen; Method: conceptual translation with partial peptide sequencing. This sequence comes from Fig. 5"
BASE COUNT		201 a	68 c 99 g 127 t
ORIGIN			
Query Match		2.9%; Score 50;	DB 36; Length 495;
Best Local Similarity		100.0%;	Pred. No. 3.3e-15;
Matches	50;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1672	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCCGC 1721	
Dn	446	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCCGC 495	
RESULT 39			
LOCUS	AF139986	3362 bp	mRNA VMT 30-MAY-1999
DEFINITION	Xenopus laevis cleavage and polyadenylation specificity factor 100 kDa subunit mRNA, complete cds.		
ACCESSION	AF139986		
NID	94927239		
VERSION	AF139986.1	GI:4927239	
KEYWORDS	African clawed frog.		
SOURCE ORGANISM	Xenopus laevis Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.		
REFERENCE	1 (bases 1 to 3362) Dickson,K.S., Bilger,A., Ballantyne,S. and Wickers,M.P. CPSP in Xenopus laevis oocytes: a cytoplasmic factor involved in regulated polyadenylation unpublished		
JOURNAL	2 (bases 1 to 3362) Dickson,K.S. and Wickers,M.P. Direct Submission Submitted (01-APR-1999) Biochemistry, UW-Madison, 433 Babcock Dr., Madison, WI 53706, USA		
FEATURES	Source		
	location/Qualifiers		
	1..3362		
	/organism="Xenopus laevis"		
	/db_xref="taxon:8355"		
	143..2494		
	/note="CPSP"		
	/codon_start=1		
	/product="cleavage and polyadenylation specificity factor 100 kDa subunit"		
	/protein_id="AAD33061.1"		
	/db_xref="PID:94927240"		
	/db_xref="GI:4927240"		
	/translation="MTSTIKITLVGAQESANCYLIVDFFRLDCGDNDFSHDI ISBVKYIHOVDVAVLSHPDPVLGLALPYAAGKLGNCATIAVIPYKMGOMPTDY QSRHNEDSLFSDLDVDCAPFKIQQLKNOIVHLKGKHGSLTPLPGHMIGTIW KLNVDEELIYVAVFNENHREIRLNGCSLEMINRPULLTDSNNAVYOPRRKORDQ LATNETLETARGGNVLIADVIAGRAVELCOLDDOIVRTDAGAGYSLLANLNVS VESRSOVEMSSMKRCPEDKRNPFORFLILICHSYDLAVPSPKVLAASOPLE CGSERLFIOWCODPKNSVILYRTTPGTARLIDHSEPRIDIPLRKRYKEGKL BEVEVEREKTRKAALKLEDSDSDSDVEEDIDOTISKAHMDLMKNEGSKR GSFPKAKSYIPMPAPEDRIKWDEYEGLIKPDEFLVPPELOVEDERTKLESOLINGD EPMDOKASVTPTKCVSTESMEIKAVTVIIDEGRSGDSIKKIINOMPROLIIHG		
CDS			

PPDATTDLAEACRAFGKDKIKVYTPKLEHETVDATSETHIYOVRLKDSLVSLKFCXAK
DTELAIDGVDIMRYSKYDITGVILEERELKDEGDEMVOVTOYMDASTIQAOKVKS
LFDDDKSESESEIIPLEPLPSNENYGHOSVMMNPERISDFOVLLRESIHAEFVG
GVLCNNMVAARRETGTGIGLGCCEGFEFRIRRELYEQVAY"

BASE COUNT 1086 a 587 c 750 g 939 t

Query Match 2.8%; Score 49; DB 4; Length 3362;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1665 ACTTAAAAA
Db 3286 ACTTAAAAA

RESULT 40

AF153205 2035 bp mRNA VRT 07-JUN-1999
LOCUS Gallus gallus CD44-like protein mRNA, complete cds.
DEFINITION
ACCESSION AF153205
NID 95002372
VERSION AF153205.1 GI:5002372
KEYWORDS
SOURCE

ORGANISM

Chicken.
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria;
Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 2035)
AUTHORS Zhu, S., Dong, S. and Halfter, W.
TITLE Transient expression of a CD44-like protein in the optic nerve and
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2035)
AUTHORS Zhu, S., Dong, S. and Halfter, W.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-1999) Neurobiology, University of Pittsburgh,
3300 Terrace Street, Pittsburgh, PA 15261, USA

FEATURES
source location/Qualifiers
1..2035
/organism="Gallus gallus"
/db_xref="taxon:9031"
/tissue_type="amnion"
/dev_stage="embryo"
75..1265
/codon_start=1
/product="CD44-like protein"
/protein_id="A037443.1"
/db_xref="PIR:G5002373"
/db_xref="GI:5002373"
/translation="MANFYLLATFGGLCLFCLTETQFNITCRGGVFEHKNRSL
TRAFAELICRALNSTLATLEQFERAHAGFETRGFIYGHVYIPRNPYHLCAANH
GIYKLSANTGTGDAYCYNAETETRSKACEIEIDITFISNOSEIYINDNDSRYAD
GTRHSGSSSTSGVDENLGSSTHDTTPGASIRSSPSVSGVTVYSHRPDSSGG
EKRFEPYKHVDEISTPTSDILATAADPPRRDDVQHPASTRSTISNDGCPHKGDGP
TSSPGLSTTTVTSOPETAHVPEMLITVAALLLAIVCNVNSRRGCGKKLVI
NNGKGAVEDKRETLNGDASKSOEMVHLVHKESQNDRTGACDEFLTVDETQNHQDGM
KSGV"

CDS

BASE COUNT 685 a 397 c 451 g 502 t
ORIGIN

Query Match 2.8%; Score 49; DB 4; Length 2035;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1665 ACTTAAAAA
Db 1951 ACTTAAAAA

RESULT 41

SACATATLAS
LOCUS
DEFINITION
ACCESSION
NID
VERSION
KEYWORDS
SOURCE
ORGANISM

SACATATLAS 1772 bp mRNA PLN 30-SEP-1997
Soldanella alpina mRNA for catalase.
299633
92462660
299633.1 GI:2462660
catalase.
Soldanella alpina.

REFERENCE 1 (bases 1 to 1772)
AUTHORS Schmidt, M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1772)
AUTHORS Schmidt, M.
JOURNAL Direct Submission
TITLE Submitted (26-SEP-1997) Schmidt M., University Frankfurt/Main,
Dept. of Botany, Stesemayerstr. 70, Frankfurt/Main, 60323, Germany
location/Qualifiers
1..1772
/organism="Soldanella alpina"
/db_xref="taxon:66308"
/clone="Cat1"
/dev_stage="Adult"
/tissue_type="leaf"
44..1522
/EC_number="1.11.1.6"
/product="catalase"
/codon_start=1
/product_id="CAB16749.1"
/db_xref="PIR:G51509"
/db_xref="PIR:G2462661"
/db_xref="GI:2462661"
/db_xref="SPTREMBL:O24339"
/translation="MDPKYRSSSAFNPFWTTNSGAPYWNNSSLTVGTRPILLED
YHLVEKLANFDEREIPERYVHARGASAKGEFEVTHDISHLCADFLRAGVOTPIVR
FSTVHERGSPETIRDPGRFAVKFYTRRENPDLGNNPVPVRGMRKFPDVAHLAKP
NRSHTQEWKRITIDFSHPRESLHMTFTPLFDLGVPODYRHRGSGVNTYTLINKAK
AOYKTRHKPCTGVCCLDEDAIKYGGAHSHATVDLDSISAGNPWKFLTIQIEP
DHEDKDFEDPDLVLTWPEDDIPLMPVGRVLNKNIDNFPAENEGALFCPALIVGVY
YSDDKLLQTRIFSYADTORHRLGPNYLOLPNAPACAHNNHHEGLMFMHDEEVN
FPSRDPVYHAEHRHPILPVLTKGKRDCHIENENFKOGEYRFFAFDPDQRFVCRW
VDALSPRYTHIRSIWISYWTQAKSGOKLASLNVRPM"

CDS

BASE COUNT 516 a 383 c 377 g 496 t
ORIGIN

Query Match 2.8%; Score 49; DB 7; Length 1772;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1665 ACTTAAAAA
Db 1720 ACTTAAAAA

RESULT 42

AF069331 711 bp mRNA PLN 17-DEC-1998
LOCUS Hordeum vulgare clone BLT1-5 low temperature induced protein mRNA,
DEFINITION complete cds.
ACCESSION AF069331
NID 94027996
VERSION AF069331.1 GI:4027996
KEYWORDS
SOURCE
ORGANISM

barley.
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Hordeum.
REFERENCE 1 (bases 1 to 711)


```

translation="MAGTAAVLPSTALPLPFGGSRAGSLRRIAPVGRSGPLTIVV
OLELKPSPYSINALEPVMQSOETLEPHGKAAHRCVGDVDMQIIGTLAGLSLEIYK
SYNCGDILLPTFNNAOAIWNDDFTFQSKPRPGCKPFLVWELLERDFGSEGMARFK
NAAITQFGSGMWALCVIKAANLDVGNANVPCPTERDYCLIIETIPNANPLIWDYNPL
VVDVWEHAYLVADVQNRPRDVSIFMDLISWEAASALEAAMQAAEAQAARAREDEE
ERRKEEDEDETERDDGGDMKMYVDSDDGDEDE"

BASE COUNT      354 a      256 c      329 g      257 t      1 others
ORIGIN

Query Match      2.8% ; Score 49; DB 8; Length 1207;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1665 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
|||||
Db 1148 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1196

RESULT 44
HSM800467      HSM800467      1845 bp      mRNA      PRI      21-MAY-1999
LOCUS          Homo sapiens mRNA; CDNA DKFZ586F1322 (from clone DKFZ586F1322).
ACCESSION      AL050172
NID            g4884384
VERSION        AL050172.1 GI:4884384
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE      1 (bases 1 to 1845)
AUTHORS        Koehler,K., Beyer,K., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE          Direct Submission
JOURNAL         Submitted (21-MAY-1999) MIPs, Am Kioferspitz 18a D-82152
Martinsried, GERMANY
COMMENT        Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); sequenced by BKFZ within the cDNA
sequencing consortium of the German Genome Project. This clone is
available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heuberweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
     source
         location/Qualifiers
             1..1845
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone="DKFZ586F1322"
                 /clone_id="586 (synonym: hute1). Vector pSport1; host
DH10B; sites NotI + SalI/MluI"
                 /dev_stage="adult"
                 /tissue_type="uterus"
                 /date="1743..1748
polyA_signal
polyA_site     1762
BASE COUNT    521 a      365 c      356 g      603 t
ORIGIN

Query Match      2.8% ; Score 49; DB 10; Length 1845;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1665 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
|||||
Db 1159 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1807

RESULT 45
MM0132433      MM0132433      1403 bp      mRNA      ROD      27-MAY-1999
LOCUS          Mus musculus untranslated RNA G90.
DEFINITION     AJ132433
ACCESSION      g4914645
NID            AJ132433.1 GI:4914645
VERSION

```

KEYWORDS G90.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
AUTHORS Krause, R., Hemberger, M., Himmelbauer, H., Kalschauer, V. and Fundele, R. H.

TITLE Identification and characterization of G90, a novel mouse RNA that lacks an extensive open reading frame

JOURNAL Gene 232 (1), 35-42 (1999)

MEDLINE 99267407
REFERENCE 2 (bases 1 to 1403)
AUTHORS Fundele, R. H.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1999) Fundele R. H., Rogers, Max-Planck-Institut fuer molekulare Genetik, Ihnestrasse 73, D-14195 Berlin, GERMANY

FEATURES
SOURCE
1. 1403
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="6"
/dev_stage="adult"
/cell_type="small intestine"
1. 1357
/note="untranslated RNA G90"
1358

misc_RNA
polyA_site 392 a 336 c 367 g 306 t 2 others
BASE COUNT
ORIGIN

Query Match 2.8%; Score 49; DB 12; Length 1403;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1665 ACTTAAA 1713
|||||
Db 1355 ACTTAAA 1403

RESULT 46
MUSCALTRA 754 bp mRNA ROD 04-FEB-1999
LOCUS Mouse mRNA for caltactin.
DEFINITION D16301
ACCESSION g467515
NID D16301.1 GI:467515
VERSION caltactin.
KEYWORDS Mus musculus (sub-species: domesticus, strain: BALB/c) testis cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota: Metazoa: Chordata: Vertebrata: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
REFERENCE 1 (bases 1 to 754)
AUTHORS Ogawa, K.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-1993) to the DDBJ/EMBL/GenBank databases. Kazuo Ogawa, National Institute for Basic Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki, Aichi 444, Japan (E-mail:ogawak@nibb.ac.jp, Tel:0564-55-7504, Fax:0564-53-7400)
2 (bases 1 to 754)
Ogawa, K. and Shimizu, T.
TITLE cDNA sequence for mouse caltactin
JOURNAL Biochim. Biophys. Acta 1216 (1), 126-128 (1993)
MEDLINE 94032472
COMMENT Submitted (20-MAY-1993) to DDBJ by: Kazuo Ogawa
Department of Cell Biology
National Institute for Basic Biology
38 Nishigonaka, Myodaiji-cho
Okazaki, Aichi 444
Japan

Phone: 0564-55-7504
Email: ogawak@nibb.ac.jp
Fax: 0564-53-7400
Location/Qualifiers
1. 754
/organism="Mus musculus"
/strain="BALB/c"
/sub-species="domesticus"
/db_xref="taxon:10090"
/tissue_type="testis"
35. 553
/codon_start=1
/product="caltactin"
/protein_id="BA03806.1"
/db_xref="PID:d1004322"
/db_xref="PID:g565281"
/db_xref="GI:565281"
/translation="MASTPRKSNVSNSTYKRVGPKPELPEDEKQEVREAFDLPSDGSGTIDYKELVAMALGFPRKPEKMKKISVDGEATGKISFNFLVAMQKMAKPTKEETLKAFRLPDDDETKISFKNLKRVANLGEGLTDEIJOEMIDEADRDGDGEVNEE EFLKIMKTMLY"

CDS
BASE COUNT 272 a 134 c 186 g 162 t
ORIGIN

Query Match 2.8%; Score 49; DB 12; Length 754;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1665 ACTTAAA 1713
|||||
Db 684 ACTTAAA 732

RESULT 47
AFCAATP 887 bp DNA INV 28-JUN-1994
LOCUS A. franciscana gene for sarco/endoplasmic reticulum Ca-ATPase.
DEFINITION X76991
ACCESSION g440382
NID X76991.1 GI:440382
VERSION Ca2+-ATPase.
KEYWORDS Artemia franciscana.
SOURCE Artemia franciscana.
ORGANISM Eukaryota: Metazoa: Arthropoda: Crustacea: Branchiopoda: Anostraca: Artemiidae: Artemia.
REFERENCE 1 (bases 1 to 887)
AUTHORS Escalante, R. and Sastre, L.
TITLE Structure of Artemia franciscana sarco/endoplasmic reticulum Ca-ATPase gene
JOURNAL J. Biol. Chem. 269 (17), 13005-13012 (1994)
MEDLINE 9420387
REFERENCE 2 (bases 1 to 887)
AUTHORS Sastre, L.
TITLE Direct Submission
JOURNAL Submitted (28-DEC-1993) L. Sastre, Inst. de Invest. Biomed. del CSIC, C/ Arturo Duperier, 4, 28029 Madrid, SPAIN

FEATURES
SOURCE
1. 887
/organism="Artemia franciscana"
/db_xref="taxon:6661"
/dev_stage="20 hours developed cysts"
/tissue_type="whole organism"
/clone_lib="EMBL3-genomic"
/clone="GAFATCa-1"
766. 887
/product="Sarco/endoplasmic reticulum Ca-ATPase"
BASE COUNT 281 a 169 c 128 g 309 t
ORIGIN

Query Match 2.8%; Score 49; DB 36; Length 887;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG 1714
 Db 463 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG 511

RESULT 48

LOCUS BMOTELRPT/c 562 bp DNA INV 26-FEB-1999
 DEFINITION Bombyx mori gene, telomeric repeat sequence.
 ACCESSION D13554
 NID 9217292
 VERSION D13554.1 GI:217292
 KEYWORDS telomeric repeat.
 SOURCE Bombyx mori (strain:Kinshu x Showa, haplotype:diploid) fifth larvae fat body DNA, clone_11b:lambda EMBL3 clone:pBluescript SK+.
 PB11-HK.

ORGANISM

Bombyx mori
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Lepidoptera; Bombycoidea; Bombycidae; Bombyx.

REFERENCE

1 (bases 1 to 562)
 Okazaki, S., Tsuchida, K., Maekawa, H., Ishikawa, H. and Fujiwara, H. Identification of a pentanucleotide telomeric sequence, (TTAGG)_n, in the silkworm Bombyx mori and in other insects

JOURNAL

Mol. Cell. Biol. 13 (3), 1424-1432 (1993)
 93180787

MEDLINE

2 (bases 1 to 562)
 Okazaki, S.
 Direct Submission
 Submitted (05-NOV-1992) to the DDBJ/EMBL/GenBank databases. Satoshi Okazaki, University of Tokyo, Faculty of Science, Zoological Institute, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan
 (Tel:03-2812-2111(ex.4449), Fax:03-3816-1965)

FEATURES

source

1..562
 /organism="Bombyx mori"
 /strain="Kinshu x Showa"
 /db_xref="taxon:7091"
 /clone="pBluescript SK+", PB11-HK"
 /clone_11b="lambda EMBL3"
 /dev_stage="fifth larvae"
 /haplotype="diploid"
 /tissue_type="fat body"
 /tissue_type="telomeric repeat"
 /rpt_family="direct"
 /rpt_type="direct"
 repeat_region
 331..459
 /rpt_family="telomeric repeat"
 /rpt_type="direct"
 BASE COUNT 122 a 83 c 134 g 223 t
 ORIGIN

Query Match 2.8%; Score 49; DB 36; Length 562;

Best Local Similarity 100.0%; Pred. No. 1..15;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1665 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
 Db 533 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 485

RESULT 49

LOCUS AF013214 1032 bp mRNA MAM 04-AUG-1997
 DEFINITION Bos taurus acidic ribosomal phosphoprotein PO mRNA, partial cds.
 ACCESSION AF013214
 NID 92293576
 VERSION AF013214.1 GI:2293576
 KEYWORDS

SOURCE

Bos taurus.
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos.

REFERENCE

1 (bases 1 to 1032)
 Lileinsiek, B., Rooha, M., Umansky, V., Benner, A., Lin, Y., Ziegler, R., Nawroth, P.P. and Schirmacher, V.

Direct Submission
 Submitted (10-JUL-1997) Internal Medicine I, University Heidelberg, Im Neuenheimer Feld 324, 3.0G, R320, Heidelberg 69120, Germany

FEATURES

source

CDS

1..910
 /note="located within the 60S subunit of eukaryotic ribosomes; interacts with elongation factor 1 alpha and elongation factor 2 alpha; involved in protein synthesis"
 /codon_start=2
 /product="acidic ribosomal phosphoprotein PO"
 /protein_id="AAB65436.1"
 /db_xref="PID:92293577"
 /db_xref="GI:2293577"
 /translation="IIQLDDYPCFIVGADNVGSKOMQOIRMSLRGAIVLGNKNTM MKRAIRGHLENNPALKELPHIRGNVGFTEDELIRDMILANKVPAARAGATAP CEVTVPAQNTGLGPEKTSFEOALGITKISRGTIELISDVLKGTGDKGASATLLN MNISPSFGVIGQVEFNGSTINPEVLTDEETLHRSRLEGRNVAASYCIOGYPTV ASYPSHITNGKRYVALSVETDVTFFPLAKRKVAFLDPSAFVAALVAAAPAAPAT TAAFAVKAKESESEDEMGFGLPD"
 BASE COUNT 283 a 271 c 264 g 214 t
 ORIGIN

Query Match 2.8%; Score 48; DB 3; Length 1032;

Best Local Similarity 100.0%; Pred. No. 3..8e-15;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
 Db 975 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1022

RESULT 50

LOCUS ECCD44 1338 bp mRNA MAM 07-SEP-1993
 DEFINITION E.caballus mRNA for lymphocyte surface antigen precursor CD44.
 ACCESSION X66862
 NID 91059
 VERSION X66862.1 GI:1059
 KEYWORDS lymphocyte surface antigen.
 SOURCE Equus caballus.
 ORGANISM Equus caballus.

REFERENCE
 1 (bases 1 to 1338)
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

1 (bases 1 to 1338)
 Tavernor, A.A.

Direct Submission
 Submitted (15-JUN-1992) A.S. Tavernor, Inst of Animal Physiology & Genetics, Research, Cambridge Station Babraham Hall, Cambridge CB2 4AT, UK

REFERENCE

2 (bases 1 to 1338)
 Tavernor, A.S., Deyerson, E.V., Coadwell, W.J., Lunn, D.P., Zhang, C., Davis, W. and Butcher, G.W.

Molecular cloning of equine CD44 cDNA by a COS cell expression system

Immunogenetics 37 (6), 474-477 (1993)

JOURNAL

MEDLINE

FEATURES

1..1338
 /organism="Equus caballus"
 /isolate="MACD44"
 /db_xref="taxon:9796"
 /tissue_type="peripheral blood lymphocyte (PBL)"

sig_peptide
CDS
/cell_type="PBL 42 hour concanavalin A blasts"
/clone_lib="CDNA"
/clone="MACE044"
55..114
/product="lymphocyte surface antigen CD44"
55..1134
/codon_start=1
/product="lymphocyte surface antigen precursor CD44"
/protein_id="CAA47331.1"
/db_xref="PID:g1060"
/db_xref="GI:1060"
/db_xref="SWISS-PROT:005078"
/translation="MDKFMWRAAMGICLVPLSLAQLDNLITCRYAGVFHVENGRYSI
SRTKADLCRAENSTIPYMAQMKALNIGFECRIGIFGHVYIPPIHNSICAAANT
GVYILITSNISQIDTYCFNASAPPEEDCTSYTDLPNAFEGPITITLYNRDGTRTKKE
YRTNPEIDINPSTPADDDVSSGSSSEKSTSGGYSIFTHLPITRPTODQSSPWYSDPE
KPTPTKDRASGGRAQTTHGSETSGHSTGSGGASTTSGPIRPPQIPFWLITLASILA
LALILAVCIAVNSRRCGKKLVINNGAVDRAKASGLNGEASRSQEWVHLVKNES
SETQDOFMTADETRNLQNVDMKIGV"
115..1131
/product="lymphocyte surface antigen CD44"
853..915
/note="transmembrane region"
1249..1254
polya_site 1269..1329
BASE COUNT 390 a 381 c 330 g 237 t
ORIGIN

Query Match 2.8%; Score 48; DB 3; Length 1338;
Best Local Similarity 100.0%; Pred. No. 3.8e-15;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
|||||
Db 1266 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1313

Search completed: August 7, 1999, 21:39:11
Job time: 1141 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 1999, 02:54:36 ; Search time 46.68 Seconds

(without alignments)
9309.847 Million cell updates/sec

Title: US-09-104-063-3

Perfect score: 1737

Sequence: 1 GAATTCAGTGTGCTGGCGG.....CCGCCAGCAGCTGGAATTC 1737

Scoring table: OLIGO_NUC

Searched: 311585 seqs, 125096042 residues

Database: N_Geneseq_36:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1737	100.0	1737	1 080521	Human monocyte PF4
2	1737	100.0	1737	1 099007	Chemokine superfamily
3	1533	88.3	1737	1 029506	New platelet factor
4	1533	87.7	1944	1 X15882	CDNA encoding G-pr
5	1067	51.4	1317	1 V18357	Human PM3 seven tr
6	1016	58.5	1317	1 066179	Seven transmembran
7	249	14.3	611	1 X15883	CDNA encoding a pa
8	218	12.6	218	1 T20146	Human gene signal
9	53	3.1	3581	1 044391	Sequence of murine
10	53	3.1	1338	1 055607	Rabbit zona pelluc
11	53	3.1	87	1 V05740	Nucleotide sequenc
12	53	3.1	94	1 V05728	Nucleotide sequenc
13	53	3.1	4237	1 V61487	Human secreted pro
14	53	3.1	1338	1 V64789	Rabbit ZPC CDNA. I
15	53	3.1	2744	1 V84799	Nucleotide sequenc
16	51	2.9	61	1 078635	Human chondromodul
17	51	2.9	775	1 X26148	Human pan-s/ck-1A
18	51	2.9	1718	1 X26151	3' CDNA sequence o
19	50	2.9	495	1 T74484	Pilostulin 2 precur
20	50	2.9	201	1 V00418	3' fragment of clo
21	50	2.9	1877	1 V46370	Nucleic acid encod
22	49	2.8	56	1 052732	Sequence of oligo
23	49	2.8	55	1 052734	Sequence of oligo
24	49	2.8	2378	1 T84975	Rat glial cell der
25	49	2.8	3443	1 V31197	Secreted protein D
26	49	2.8	2519	1 V54124	Human membrane pro
27	49	2.8	1770	1 X52242	Protein PRO233 CN
28	48	2.8	1585	1 047967	Rape acyl-ACP thio
29	48	2.8	699	1 058663	Human superoxide d
30	48	2.8	1490	1 075361	Glibetrelilin-20-ox
31	48	2.8	861	1 098552	Thrombopolein cod
32	48	2.8	1474	1 T90174	Oil seed rape cyst
33	48	2.8	2312	1 V21646	T. gondii chorisma
34	48	2.8	5503	1 V18187	Fanconi anaemia of
35	48	2.8	90	1 V37789	Analytical solid p
36	48	2.8	1049	1 V43610	Human secreted pro
37	48	2.8	772	1 V58363	Coding sequence fo
38	48	2.8	1328	1 V58754	Human secreted pro
39	48	2.8	1373	1 V52969	Human G1 protein 1
40	48	2.8	1579	1 V64590	Tobacco chili DNA.
41	48	2.8	2120	1 X33812	Coding sequence fo
42	48	2.8	1090	1 X37521	Human secreted pro
43	47	2.7	1239	1 N91576	Sequence of clone

44	47	2.7	2549	1 N91467	Sequence of human
45	47	2.7	1834	1 004690	Encodes Mammalian
46	47	2.7	372	1 N80489	Cowpea trypsin inh
47	47	2.7	7722	1 Q12023	Norwalk virus geno
48	47	2.7	86	1 Q11760	Self-complementary
49	47	2.7	1793	1 Q15139	Clone pLAC6 encod
50	47	2.7	1410	1 Q15022	Hyoxyamine 6 beta
51	47	2.7	1118	1 Q0263	Deg-1 gene. DNA se
52	47	2.7	1239	1 Q23000	Sequence encoding
53	47	2.7	766	1 N40162	Sequence of prepro
54	47	2.7	1396	1 Q24371	Natural killer cel
55	47	2.7	1078	1 Q25785	Protein synthesis
56	47	2.7	2116	1 Q35987	Tomato hsp80 CDNA
57	47	2.7	1239	1 Q43439	LOI PI CDNA clone
58	47	2.7	222	1 Q46071	Sequence downstre
59	47	2.7	1453	1 Q50573	Asparaginylendopep
60	47	2.7	1640	1 Q50575	Asparaginylendopep
61	47	2.7	1910	1 Q50579	Asparaginylendopep
62	47	2.7	120	1 Q52429	Human subtelomeric
63	47	2.7	1738	1 Q56241	Sequence encoding
64	47	2.7	7753	1 Q56826	Norwalk virus stia
65	47	2.7	1172	1 Q58869	Protein allergen o
66	47	2.7	1393	1 Q66241	Alpha-galactosidas
67	47	2.7	600	1 Q72739	T. ovis vaccine ca
68	47	2.7	1393	1 Q77861	Human alpha-galact
69	47	2.7	458	1 Q90525	Rat Sirt 15 kDa su
70	47	2.7	1748	1 Q88327	Arabidopsis thalia
71	47	2.7	120	1 Q87011	Subtelomeric CDNA
72	47	2.7	399	1 T13044	Cotton fibre-speci
73	47	2.7	767	1 T15821	TBE virus strain N
74	47	2.7	1141	1 T15820	TBE virus strain N
75	47	2.7	600	1 T17934	Taenia ovis anti
76	47	2.7	1582	1 T16831	Human survival mot
77	47	2.7	1560	1 T18828	Human survival mot
78	47	2.7	1558	1 T18255	Human survival mot
79	47	2.7	1582	1 T28259	Survival motor neu
80	47	2.7	399	1 T30261	Cotton fibre cell-
81	47	2.7	1141	1 T44469	Cotton fibre cell-
82	47	2.7	767	1 T44470	Tick-borne enceph
83	47	2.7	1364	1 T48099	Human interleukin-
84	47	2.7	3524	1 T46186	CDNA encoding cell
85	47	2.7	1174	1 T39050	Cotton fibre speci
86	47	2.7	399	1 T62620	CDNA encoding anti
87	47	2.7	916	1 T65651	Rat RNA polymerase
88	47	2.7	458	1 T59993	PTH-like peptide D
89	47	2.7	2123	1 T59701	Human origin of re
90	47	2.7	3214	1 T63260	Mouse CD100 antige
91	47	2.7	4391	1 T60665	Human c-IAP1. Nuci
92	47	2.7	2589	1 T61590	Cotton fibre speci
93	47	2.7	399	1 T70051	Poplar 1-aminocycl
94	47	2.7	1868	1 T73502	Human origin of re
95	47	2.7	3214	1 T73287	Clone H174-10. Pol
96	47	2.7	966	1 T87430	Subtelomeric CDNA
97	47	2.7	120	1 T86338	Human prostate pro
98	47	2.7	375	1 T84939	Human secreted pro
99	47	2.7	69	1 V02148	3' fragment of clo
100	47	2.7	101	1 V00420	Human M97-2 secret
101	47	2.7	99	1 T91300	Human chemokine re
102	47	2.7	1316	1 T98543	3' portion of CDNA
103	47	2.7	69	1 T88081	Wheat soluble stia
104	47	2.7	2239	1 V01527	Nucleotide sequenc
105	47	2.7	85	1 V05720	Wild tomato acyltr
106	47	2.7	1604	1 V04252	Tomato ACC synthas
107	47	2.7	2230	1 V15704	Heterodimeric huma
108	47	2.7	1364	1 V15228	Hom sapiens clone
109	47	2.7	144	1 V21239	Mus musculus Tub I
110	47	2.7	1700	1 V11858	Oligonucleotide se
111	47	2.7	80	1 V37197	Human novel secret
112	47	2.7	79	1 V26706	Monocomponent endo
113	47	2.7	1174	1 V39096	Nucleotide sequenc
114	47	2.7	4358	1 V42653	Human interleukin-
115	47	2.7	79	1 V32414	
116	47	2.7	1364	1 V42536	

117	47	2.7	568	1	V53256	ETH precursor anti
118	47	2.7	2790	1	V32998	Human concentrativ
119	47	2.7	1508	1	V07665	Mauke Bx1 cDNA (DI
120	47	2.7	1694	1	V33190	Secreted protein B
121	47	2.7	1753	1	V59556	Human secreted pro
122	47	2.7	2682	1	V63195	CDNA from clone 9g
123	47	2.7	916	1	V62000	R. prolixus NO-r p
124	47	2.7	1406	1	V61486	Human secreted pro
125	47	2.7	1114	1	V34315	Human secreted pro
126	47	2.7	1123	1	V34290	Human secreted pro
127	47	2.7	1147	1	V69620	Human secreted full len
128	47	2.7	568	1	V34054	Porcine reproductive
129	47	2.7	98	1	X00174	Human secreted pro
130	47	2.7	1376	1	X04325	Human adult trache
131	47	2.7	1240	1	X06787	Human sapiens fetal
132	47	2.7	1230	1	X07565	Human secreted pro
133	47	2.7	826	1	X27402	Human secreted pro
134	47	2.7	1223	1	X25130	Soybean isoflavone
135	47	2.7	997	1	X52274	Protein PRO244 cDN
136	47	2.7	546	1	X37513	CDNA sequence enco
137	46	2.6	2676	1	002819	Papaya ringspot vi
138	46	2.6	1124	1	003669	Sequence of plasmid
139	46	2.6	3926	1	003735	Plasmid pfgHl enco
140	46	2.6	3820	1	003517	Cucumber mosaic vi
141	46	2.6	1423	1	N90249	DNA sequence of ri
142	46	2.6	1923	1	N91039	Human pro-urokinas
143	46	2.6	2427	1	004107	Aequorin gene Bios
144	46	2.6	957	1	004441	Entire porcine tra
145	46	2.6	2671	1	003303	Clone lmd2 encodin
146	46	2.6	3504	1	005304	Sequence of new pl
147	46	2.6	3850	1	N81634	pMO440 aequorin ge
148	46	2.6	959	1	N81534	Sequence encoding
149	46	2.6	3784	1	005336	Plasminogen gene f
150	46	2.6	6020	1	006648	Sequence of CDNA 1
151	46	2.6	3905	1	N70461	Sequence of CDNA 1
152	46	2.6	1801	1	N70524	Sequence of CDNA 1
153	46	2.6	1922	1	N70525	DNA encoding human
154	46	2.6	2035	1	N70687	Encodes Xenopus Bo
155	46	2.6	1558	1	010856	Plasmid pMG4B12 us
156	46	2.6	748	1	010376	Self-complementary
157	46	2.6	90	1	011762	Human plasminogen
158	46	2.6	6010	1	011998	Fragment D of urac
159	46	2.6	1013	1	012431	Human pro-growth h
160	46	2.6	3557	1	N60801	Plasmid sequence e
161	46	2.6	3547	1	N60846	Plasmid sequence e
162	46	2.6	3628	1	N60848	Human pre-prolacti
163	46	2.6	3621	1	N60847	Rhizopus PGK1 prom
164	46	2.6	2051	1	Q14944	Human pro-urokinas
165	46	2.6	2377	1	Q20360	Sequence of a modif
166	46	2.6	941	1	N30062	Human Factor XIII
167	46	2.6	3905	1	Q25893	T23 cDNA. Stamen-s
168	46	2.6	347	1	Q27482	CA455. Anthr-spec
169	46	2.6	796	1	Q27948	Korean hepatitis C
170	46	2.6	9472	1	Q33282	NANB virus strain
171	46	2.6	9589	1	Q38218	Rape acyl-ACP thio
172	46	2.6	1710	1	Q47966	Human anti-HBs lig
173	46	2.6	1066	1	Q49943	Sequence of the 3'
174	46	2.6	289	1	Q53450	Human anti-HBs lig
175	46	2.6	1041	1	Q54651	T84.12 light chain
176	46	2.6	1041	1	Q54653	T84.12 L4-12-1 lig
177	46	2.6	2669	1	Q56925	Pig TGF-beta-3. Nu
178	46	2.6	667	1	Q61566	Human SOD. Recombi
179	46	2.6	688	1	Q61572	Human SOD. Recombi
180	46	2.6	667	1	Q61575	Hybrid human/pig S
181	46	2.6	5994	1	Q65674	Sequence encoding
182	46	2.6	2405	1	Q70827	Protein kinase (CK
183	46	2.6	3344	1	Q80228	Rat NDF clone 22 D
184	46	2.6	2335	1	Q80216	Human PRNDF-alpha
185	46	2.6	1651	1	Q80218	Human NDF-alpha2b
186	46	2.6	2760	1	Q85425	T4 DNA-polymerase
187	46	2.6	2760	1	Q85426	T4 DNA-polymerase
188	46	2.6	826	1	Q79736	Flower style-speci
189	46	2.6	826	1	Q79736	Flower style-speci
190	46	2.6	779	1	Q79737	Flower style-speci
191	46	2.6	2430	1	Q83161	DNA encoding malic
192	46	2.6	1813	1	Q88760	Human ubiquitous n
193	46	2.6	1536	1	Q94111	mML genomic DNA. T
194	46	2.6	2405	1	Q92962	Human HRR25-like c
195	46	2.6	1461	1	T01471	Mammalian mast cel
196	46	2.6	141	1	T06012	Immunodominant tra
197	46	2.6	570	1	T13041	Cotton fibre-speci
198	46	2.6	1641	1	T13041	Human interleukin-
199	46	2.6	1534	1	T15758	OR-1 orphan recept
200	46	2.6	1334	1	T15229	Spinach debranchin
201	46	2.6	3437	1	T31478	Kinetochore protei
202	46	2.6	10136	1	T34578	Rat vas deferens p
203	46	2.6	1837	1	T33852	Human p57 coding s
204	46	2.6	191	1	T13475	Capture probe for
205	46	2.6	570	1	T30258	Cotton fibre cell-
206	46	2.6	350	1	T18200	Infectious bursal
207	46	2.6	1154	1	T39048	CDNA encoding cell
208	46	2.6	826	1	T64554	Tomato S-ribonucle
209	46	2.6	609	1	T64554	Tomato S-ribonucle
210	46	2.6	903	1	T62617	Cotton fibre speci
211	46	2.6	1166	1	T60872	Cotton fibre speci
212	46	2.6	570	1	T70048	Mouse lipopolysacc
213	46	2.6	2400	1	T49364	Cotton fibre speci
214	46	2.6	1113	1	T75765	Human g protein ga
215	46	2.6	13414	1	T71321	Full length 3' UTR
216	46	2.6	240	1	T76782	Plasmid pC851 enco
217	46	2.6	140	1	T76781	DNA encoding human
218	46	2.6	1001	1	T73851	Cucumber mosaic vi
219	46	2.6	10288	1	T71522	Yellowtail tuna DN
220	46	2.6	1426	1	T72272	Yellow tail/fin tu
221	46	2.6	1898	1	T79634	Mouse receptor ME2
222	46	2.6	3851	1	T79857	Human NB phox 3' u
223	46	2.6	851	1	T90019	Alzheimer's diseas
224	46	2.6	10195	1	T85320	Alzheimer's diseas
225	46	2.6	6791	1	T85319	Alzheimer's diseas
226	46	2.6	1769	1	T85973	Alzheimer's diseas
227	46	2.6	1420	1	T72167	Alzheimer's diseas
228	46	2.6	1858	1	T72175	Alzheimer's diseas
229	46	2.6	3527	1	T72171	Human immunodefici
230	46	2.6	5173	1	T89783	Malassezia fungus
231	46	2.6	782	1	T85876	Malassezia fungus
232	46	2.6	818	1	T85877	Malassezia fungus
233	46	2.6	1776	1	T91744	Tobacco calcium/ca
234	46	2.6	1577	1	T90169	Al seed rape cyst
235	46	2.6	4586	1	T96838	Intron 8 of human
236	46	2.6	259	1	V00423	3' fragment of clo
237	46	2.6	3451	1	V02308	Cell membrane prot
238	46	2.6	615	1	V02861	Human HMG1-C aberr
239	46	2.6	216	1	V09114	3' nucleotide sequ
240	46	2.6	2205	1	V04680	Human presenilin 1
241	46	2.6	741	1	V23153	3' nucleotide port
242	46	2.6	740	1	V23878	Plant CCR enzyme D
243	46	2.6	1670	1	V23915	Plant CCR enzyme D
244	46	2.6	1844	1	V26557	Human IP-10/Mig re
245	46	2.6	4931	1	V24018	Human BHL1 coding
246	46	2.6	1711	1	V29584	Homo sapiens sulph
247	46	2.6	1728	1	V33617	Human secreted pro
248	46	2.6	2836	1	V40744	Human receptor for
249	46	2.6	2836	1	V40745	C. felis esterase,
250	46	2.6	2307	1	V47142	Novel haemopoietin
251	46	2.6	4329	1	V44318	Novel starch-asso
252	46	2.6	2026	1	V46313	Seq ID #7 from DEL
253	46	2.6	2837	1	V46317	Human secreted pro
254	46	2.6	921	1	V40521	Human secreted pro
255	46	2.6	1505	1	V40522	Homo sapiens CH27
256	46	2.6	2214	1	V40524	Homo sapiens C1542
257	46	2.6	1332	1	V40647	Homo sapiens C0122
258	46	2.6	1627	1	V40646	Human transaldolas
259	46	2.6	1627	1	V40646	Human secreted pro
260	46	2.6	1627	1	V40646	Human secreted pro
261	46	2.6	1627	1	V40646	Human secreted pro
262	46	2.6	1627	1	V40646	Human secreted pro

263	46	2	6	2892	1	V58361	Coding sequence fo
264	46	2	6	2280	1	V52504	Human metalloprote
265	46	2	6	2639	1	V52934	Pig transforming g
266	46	2	6	1836	1	V63461	Human T1-receptor
267	46	2	6	3899	1	V33929	cDNA from clone dt
268	45	2	6	577	1	V59806	Human secreted pro
269	46	2	6	1378	1	V59706	Human secreted pro
270	46	2	6	1509	1	V59679	Human secreted pro
271	46	2	6	2196	1	V59663	Human secreted pro
272	46	2	6	1143	1	V59619	Human secreted pro
273	46	2	6	975	1	V61499	Bak binding protei
274	46	2	6	1029	1	V45444	Human chemokine zs
275	46	2	6	1599	1	V34294	Human secreted pro
276	46	2	6	100	1	V68827	DNA molecule encod
277	46	2	6	4051	1	V62572	Tumour necrosis fa
278	46	2	6	3964	1	V62673	Tumour necrosis fa
279	46	2	6	340	1	V89136	EST clone BR309. N
280	46	2	6	2271	1	V84632	Human secreted pro
281	46	2	6	2276	1	V84583	Human secreted pro
282	46	2	6	2218	1	V84503	Human secreted pro
283	46	2	6	2483	1	V84468	Human secreted pro
284	46	2	6	1686	1	V73012	Human adult brain
285	46	2	6	2294	1	V79584	Rat organic anion
286	46	2	6	1733	1	V81394	Human tumour anti
287	46	2	6	1015	1	X00713	Human secreted pro
288	46	2	6	1020	1	X00681	Human secreted pro
289	46	2	6	1215	1	X00620	Human secreted pro
290	46	2	6	1153	1	V55748	Human secreted pro
291	46	2	6	3213	1	V55742	Human secreted pro
292	46	2	6	645	1	X04348	Human secreted pro
293	46	2	6	1261	1	X04382	Human secreted pro
294	46	2	6	890	1	X04376	Human secreted pro
295	46	2	6	722	1	X06786	Human adult uterus
296	46	2	6	3059	1	V99912	Human salivadin
297	46	2	6	3099	1	V99911	Human salivadin
298	46	2	6	1146	1	X07000	Xenopus Frazzled p
299	46	2	6	2264	1	X20496	Human secreted pro
300	46	2	6	2265	1	X20435	Human secreted pro
301	46	2	6	602	1	X20419	Human secreted pro
302	46	2	6	2888	1	X22123	Human secreted pro
303	46	2	6	1332	1	X02709	Human transaldol
304	46	2	6	832	1	X30363	DNA encoding a hum
305	46	2	6	1230	1	X22276	Human secreted pro
306	46	2	6	5267	1	X21357	Human BAI3 gene. N
307	46	2	6	1564	1	X22331	Human secreted pro
308	46	2	6	1654	1	X07431	Homo sapiens secre
309	46	2	6	1835	1	X27245	Human CLAR1 codin
310	46	2	6	3375	1	X28358	Human Stat6 codin
311	46	2	6	886	1	X27366	Human secreted pro
312	46	2	6	865	1	X51754	DNA encoding a hum
313	46	2	6	1361	1	X51714	DNA encoding a hum
314	46	2	6	487	1	X30156	Human secreted pro
315	46	2	6	935	1	X51702	DNA encoding a hum
316	46	2	6	1367	1	X51727	DNA encoding a hum
317	46	2	6	1549	1	X24921	Guanine ras carboxy-
318	46	2	6	650	1	X37406	Human secreted pro
319	46	2	6	824	1	X37387	Human secreted pro
320	46	2	6	773	1	X37388	Human secreted pro
321	46	2	6	604	1	X37373	Human secreted pro
322	46	2	6	2487	1	X33813	Coding sequence fo
323	46	2	6	7797	1	X33180	Cowpox virus bsr f
324	46	2	6	6644	1	X33181	Base sequence of t
325	46	2	6	7372	1	X33182	Base sequence of t
326	46	2	6	7996	1	X33184	Base sequence of t
327	46	2	6	980	1	X37465	Human secreted pro
328	46	2	6	971	1	X37469	Human secreted pro
329	46	2	6	1510	1	X03740	Human myb related
330	46	2	6	669	1	N90103	Human prealbumin c
331	45	2	6	4180	1	N91773	Rat androgen recep
332	45	2	6	1973	1	N91045	Gene encoding bovl
333	45	2	6	1493	1	O04492	Sequence encoding
334	45	2	6	960	1	O03293	Recombinant DNA en
335	45	2	6	2400	1	O05055	Placenta-specific

409	45	2.6	1623	1	114925	T cell replacing f
410	45	2.6	2158	1	115228	Tumor necrosis fac
411	45	2.6	1721	1	138129	Nuclear envelope-a
412	45	2.6	1425	1	135220	Cytoplasmic antipr
413	45	2.6	1016	1	129401	Lipase coding sequ
414	45	2.6	141	1	134595	Probe for detectin
415	45	2.6	635	1	138390	Murine glycosylati
416	45	2.6	630	1	145865	CDNA encoding C. b
417	45	2.6	1602	1	143204	Human microg-en-acti
418	45	2.6	1964	1	145937	Carcinoma-actin a
419	45	2.6	3736	1	145937	Human hypoxia indu
420	45	2.6	940	1	158281	Arabidopsis Stz po
421	45	2.6	1276	1	142071	Human nuclear prol
422	45	2.6	3040	1	142853	Phospholipase D pr
423	45	2.6	780	1	159912	Human transcriptio
424	45	2.6	2001	1	159700	PTH-like peptide D
425	45	2.6	3871	1	169592	Human Ob receptor
426	45	2.6	1849	1	174042	Soybean thiol prot
427	45	2.6	90	1	174082	Synthetic DNA prob
428	45	2.6	90	1	174081	Synthetic DNA prob
429	45	2.6	5894	1	185474	hABC3 CDNA sequenc
430	45	2.6	1713	1	185993	Maize 5-enolpyruv
431	45	2.6	1713	1	186000	Maize 5-enolpyruv
432	45	2.6	558	1	164553	Mouse thymus and a
433	45	2.6	204	1	161958	Detection probe fo
434	45	2.6	249	1	181968	Hepatitis C virus
435	45	2.6	227	1	189600	Hepatitis C virus
436	45	2.6	270	1	189602	Hepatitis C virus
437	45	2.6	260	1	189601	Hepatitis C virus
438	45	2.6	356	1	189597	Hepatitis C virus
439	45	2.6	257	1	189589	Hepatitis C virus
440	45	2.6	1023	1	170132	Max-interacting pr
441	45	2.6	253	1	185891	Alzheimer's diseas
442	45	2.6	1512	1	172173	Alzheimer's diseas
443	45	2.6	882	1	172172	Human double stran
444	45	2.6	6671	1	189709	Phospholipase D en
445	45	2.6	3040	1	185509	TNF-RI-DI ligand p
446	45	2.6	2158	1	194631	Human PKR gene. Sc
447	45	2.6	2628	1	191060	Rat neurotrophin y
448	45	2.6	2481	1	187940	Rat neurotrophin y
449	45	2.6	2604	1	173595	Human RHAM genom
450	45	2.6	1040	1	1702818	Human RHAM CDNA.
451	45	2.6	3114	1	190173	Oil seed rape cyst
452	45	2.6	1441	1	190173	3' fragment of clo
453	45	2.6	106	1	190430	Human cytoplasmic
454	45	2.6	730	1	190430	Human RNA-binding
455	45	2.6	863	1	191302	CDNA for wild type
456	45	2.6	877	1	191302	Human retinoid rec
457	45	2.6	5894	1	191302	CDNA encoding full
458	45	2.6	1817	1	191302	3' nucleotide sequ
459	45	2.6	2096	1	191302	Rat FRG2 coding s
460	45	2.6	2233	1	191302	Human cyclin DI on
461	45	2.6	6525	1	191302	Secreted protein C
462	45	2.6	208	1	191302	Nucleotide sequenc
463	45	2.6	1958	1	191302	Pea plastidial pho
464	45	2.6	1780	1	191302	Chimeric protease
465	45	2.6	1325	1	191302	Human haematopoiet
466	45	2.6	84	1	191302	Equine arteritis v
467	45	2.6	84	1	191302	Human p26 CDNA seq
468	45	2.6	2182	1	191302	Human p26 CDNA seq
469	45	2.6	11326	1	191302	Human p26 CDNA seq
470	45	2.6	11326	1	191302	Human p26 CDNA seq
471	45	2.6	12827	1	191302	Human p26 CDNA seq
472	45	2.6	15528	1	191302	Human p26 CDNA seq
473	45	2.6	889	1	191302	Human p26 CDNA seq
474	45	2.6	69	1	191302	Human p26 CDNA seq
475	45	2.6	69	1	191302	Human p26 CDNA seq
476	45	2.6	117	1	191302	Human p26 CDNA seq
477	45	2.6	69	1	191302	Human p26 CDNA seq
478	45	2.6	85	1	191302	Human p26 CDNA seq
479	45	2.6	1454	1	191302	Human p26 CDNA seq
480	45	2.6	1474	1	191302	Human p26 CDNA seq
481	45	2.6	1749	1	191302	Human p26 CDNA seq
482	45	2.6	2065	1	191302	Human p26 CDNA seq
483	45	2.6	1525	1	191302	Human p26 CDNA seq
484	45	2.6	90	1	191302	Human p26 CDNA seq
485	45	2.6	1386	1	191302	Human p26 CDNA seq
486	45	2.6	127	1	191302	Human p26 CDNA seq
487	45	2.6	92	1	191302	Human p26 CDNA seq
488	45	2.6	106	1	191302	Human p26 CDNA seq
489	45	2.6	80	1	191302	Human p26 CDNA seq
490	45	2.6	70	1	191302	Human p26 CDNA seq
491	45	2.6	3441	1	191302	Human p26 CDNA seq
492	45	2.6	873	1	191302	Human p26 CDNA seq
493	45	2.6	2655	1	191302	Human p26 CDNA seq
494	45	2.6	553	1	191302	Human p26 CDNA seq
495	45	2.6	625	1	191302	Human p26 CDNA seq
496	45	2.6	6671	1	191302	Human p26 CDNA seq
497	45	2.6	1083	1	191302	Human p26 CDNA seq
498	45	2.6	1494	1	191302	Human p26 CDNA seq
499	45	2.6	112	1	191302	Human p26 CDNA seq
500	45	2.6	1761	1	191302	Human p26 CDNA seq
501	45	2.6	202	1	191302	Human p26 CDNA seq
502	45	2.6	525	1	191302	Human p26 CDNA seq
503	45	2.6	85	1	191302	Human p26 CDNA seq
504	45	2.6	479	1	191302	Human p26 CDNA seq
505	45	2.6	1252	1	191302	Human p26 CDNA seq
506	45	2.6	1910	1	191302	Human p26 CDNA seq
507	45	2.6	1232	1	191302	Human p26 CDNA seq
508	45	2.6	119	1	191302	Human p26 CDNA seq
509	45	2.6	1097	1	191302	Human p26 CDNA seq
510	45	2.6	2667	1	191302	Human p26 CDNA seq
511	45	2.6	2082	1	191302	Human p26 CDNA seq
512	45	2.6	1642	1	191302	Human p26 CDNA seq
513	45	2.6	1310	1	191302	Human p26 CDNA seq
514	45	2.6	719	1	191302	Human p26 CDNA seq
515	45	2.6	3395	1	191302	Human p26 CDNA seq
516	45	2.6	4874	1	191302	Human p26 CDNA seq
517	45	2.6	2447	1	191302	Human p26 CDNA seq
518	45	2.6	1713	1	191302	Human p26 CDNA seq
519	45	2.6	550	1	191302	Human p26 CDNA seq
520	45	2.6	1621	1	191302	Human p26 CDNA seq
521	45	2.6	3410	1	191302	Human p26 CDNA seq
522	45	2.6	618	1	191302	Human p26 CDNA seq
523	45	2.6	509	1	191302	Human p26 CDNA seq
524	45	2.6	470	1	191302	Human p26 CDNA seq
525	45	2.6	789	1	191302	Human p26 CDNA seq
526	45	2.6	1313	1	191302	Human p26 CDNA seq
527	45	2.6	1621	1	191302	Human p26 CDNA seq
528	45	2.6	3410	1	191302	Human p26 CDNA seq
529	45	2.6	732	1	191302	Human p26 CDNA seq
530	45	2.6	1153	1	191302	Human p26 CDNA seq
531	45	2.6	2323	1	191302	Human p26 CDNA seq
532	45	2.6	509	1	191302	Human p26 CDNA seq
533	45	2.6	470	1	191302	Human p26 CDNA seq
534	45	2.6	789	1	191302	Human p26 CDNA seq
535	45	2.6	2496	1	191302	Human p26 CDNA seq
536	45	2.6	2764	1	191302	Human p26 CDNA seq
537	45	2.6	819	1	191302	Human p26 CDNA seq
538	45	2.6	1307	1	191302	Human p26 CDNA seq
539	45	2.6	1772	1	191302	Human p26 CDNA seq
540	45	2.6	2025	1	191302	Human p26 CDNA seq
541	45	2.6	1181	1	191302	Human p26 CDNA seq
542	45	2.6	1212	1	191302	Human p26 CDNA seq
543	45	2.6	1605	1	191302	Human p26 CDNA seq
544	45	2.6	1864	1	191302	Human p26 CDNA seq
545	45	2.6	1864	1	191302	Human p26 CDNA seq
546	45	2.6	12980	1	191302	Human p26 CDNA seq
547	45	2.6	92	1	191302	Human p26 CDNA seq
548	45	2.6	2045	1	191302	Human p26 CDNA seq
549	45	2.6	700	1	191302	Human p26 CDNA seq
550	45	2.6	737	1	191302	Human p26 CDNA seq
551	45	2.6	1925	1	191302	Human p26 CDNA seq
552	45	2.6	1688	1	191302	Human p26 CDNA seq
553	45	2.6	1688	1	191302	Human p26 CDNA seq
554	45	2.6	1623	1	191302	Human p26 CDNA seq
555	45	2.6	1623	1	191302	Human p26 CDNA seq
556	45	2.6	1623	1	191302	Human p26 CDNA seq
557	45	2.6	1623	1	191302	Human p26 CDNA seq
558	45	2.6	1623	1	191302	Human p26 CDNA seq
559	45	2.6	1623	1	191302	Human p26 CDNA seq
560	45	2.6	1623	1	191302	Human p26 CDNA seq
561	45	2.6	1623	1	191302	Human p26 CDNA seq
562	45	2.6	1623	1	191302	Human p26 CDNA seq
563	45	2.6	1623	1	191302	Human p26 CDNA seq
564	45	2.6	1623	1	191302	Human p26 CDNA seq
565	45	2.6	1623	1	191302	Human p26 CDNA seq
566	45	2.6	1623	1	191302	Human p26 CDNA seq
567	45	2.6	1623	1	191302	Human p26 CDNA seq
568	45	2.6	1623	1	191302	Human p26 CDNA seq
569	45	2.6	1623	1	191302	Human p26 CDNA seq
570	45	2.6	1623	1	191302	Human p26 CDNA seq
571	45	2.6	1623	1	191302	Human p26 CDNA seq
572	45	2.6	1623	1	191302	Human p26 CDNA seq
573	45	2.6	1623	1	191302	Human p26 CDNA seq
574	45	2.6	1623	1	191302	Human p26 CDNA seq
575	45	2.6	1623	1	191302	Human p26 CDNA seq
576	45	2.6	1623	1	191302	Human p26 CDNA seq
577	45	2.6	1623	1	191302	Human p26 CDNA seq
578	45	2.6	1623	1	191302	Human p26 CDNA seq
579	45	2.6	1623	1	191302	Human p26 CDNA seq
580	45	2.6	1623	1	191302	Human p26 CDNA seq
581	45	2.6	1623	1	191302	Human p26 CDNA seq
582	45	2.6	1623	1	191302	Human p26 CDNA seq
583	45	2.6	1623	1	191302	Human p26 CDNA seq
584	45	2.6	1623	1	191302	Human p26 CDNA seq
585	45	2.6	1623	1	191302	Human p26 CDNA seq
586	45	2.6	1623	1	191302	Human p26 CDNA seq
587	45	2.6	1623	1	191302	Human p26 CDNA seq
588	45	2.6	1623	1	191302	Human p26 CDNA seq
589	45	2.6	1623	1	191302	Human p26 CDNA seq
590	45	2.6	1623	1	191302	Human p26 CDNA seq
591	45	2.6	1623	1	191302	Human p26 CDNA seq
592	45	2.6	1623	1	191302	Human p26 CDNA seq
593	45	2.6	1623	1	191302	Human p26 CDNA seq
594	45	2.6	1623	1	191302	Human p26 CDNA seq
595	45	2.6	1623	1	191302	Human p26 CDNA seq
596	45	2.6	1623	1	191302	Human p26 CDNA seq

555	45	2.6	1717	1	V08170	MMP19 coding seque
556	45	2.6	1152	1	V62754	Human secreted pro
557	45	2.6	378	1	V90045	EST clone CM150.
558	45	2.6	283	1	V89892	EST clone CW762. N
559	45	2.6	2010	1	V08391	Human Parathyroid
560	45	2.6	579	1	V82799	ATG-1100 (alligraf
561	45	2.6	1181	1	V81616	Cucumber peroxidase
562	45	2.6	2017	1	V08830	Gene No. 20 encodi
563	45	2.6	1705	1	V08856	Gene encoding huma
564	45	2.6	1705	1	V84516	Human secreted pro
565	45	2.6	941	1	V84586	Human secreted pro
566	45	2.6	1358	1	V84593	Human secreted pro
567	45	2.6	1637	1	V84578	Human secreted pro
568	45	2.6	472	1	V84530	Human secreted pro
569	45	2.6	1059	1	V84474	Human secreted pro
570	45	2.6	588	1	V84464	Human secreted pro
571	45	2.6	1210	1	V84425	Human secreted pro
572	45	2.6	467	1	V64423	Human secreted pro
573	45	2.6	77	1	V77194	Mouse developing 1
574	45	2.6	1521	1	V73000	Staphylococcus aur
575	45	2.6	1308	1	V73002	Human foetal brain
576	45	2.6	3076	1	V73003	Human adult brain
577	45	2.6	2171	1	V73006	Human adult brain
578	45	2.6	1181	1	V62867	Cucumber peroxidase
579	45	2.6	1456	1	V81748	Human ALP encoding
580	45	2.6	1525	1	X00679	Human secreted pro
581	45	2.6	1668	1	X00664	Human secreted pro
582	45	2.6	1507	1	X00656	Human secreted pro
583	45	2.6	2788	1	V55745	Human secreted pro
584	45	2.6	1813	1	V70895	CDNA encoding an a
585	45	2.6	1125	1	V80740	Human secreted pro
586	45	2.6	1337	1	X04345	Human secreted pro
587	45	2.6	667	1	X04334	Human secreted pro
588	45	2.6	2541	1	X04343	Human secreted pro
589	45	2.6	1322	1	X04402	Human secreted pro
590	45	2.6	2298	1	X06781	Human testis secre
591	45	2.6	1180	1	V84347	Human Apo-2Dcr CDN
592	45	2.6	74	1	V99725	Human adult testis
593	45	2.6	1554	1	X16675	Xenopus MA545 prot
594	45	2.6	2880	1	X18193	hNRC3a coding sequ
595	45	2.6	1602	1	X07066	Human mitogen acti
596	45	2.6	3812	1	X18192	hNRC3 coding sequ
597	45	2.6	1761	1	X20412	Human secreted pro
598	45	2.6	1071	1	X20414	Human secreted pro
599	45	2.6	4765	1	X09003	Mouse semaphorin r
600	45	2.6	4780	1	X09004	Mouse semaphorin r
601	45	2.6	687	1	X30410	DNA encoding a hum
602	45	2.6	1095	1	X30366	DNA encoding a hum
603	45	2.6	687	1	X30351	DNA encoding a hum
604	45	2.6	513	1	X22214	Human secreted pro
605	45	2.6	492	1	X22223	Human secreted pro
606	45	2.6	1613	1	X26196	DNA sequence of q1
607	45	2.6	2634	1	X27062	S. tuberculosis isom
608	45	2.6	3684	1	X21560	Human N-arginine d
609	45	2.6	2152	1	X22002	Human MED1 endonuc
610	45	2.6	2345	1	X18951	Human PIGR-1 encod
611	45	2.6	337	1	X25860	Clone Pegen-42 seq
612	45	2.6	588	1	X07427	Homo sapiens secre
613	45	2.6	2773	1	X07428	Homo sapiens secre
614	45	2.6	791	1	X19493	Human secreted pro
615	45	2.6	5668	1	X27262	Prostate-tumour de
616	45	2.6	3648	1	X27277	Prostate-tumour de
617	45	2.6	1129	1	X27340	Human secreted pro
618	45	2.6	921	1	X27330	Human secreted pro
619	45	2.6	553	1	X51720	DNA encoding a hum
620	45	2.6	1032	1	X30155	Human secreted pro
621	45	2.6	340	1	X51737	DNA encoding a hum
622	45	2.6	1442	1	X51744	DNA encoding a hum
623	45	2.6	990	1	X51748	DNA encoding a hum
624	45	2.6	978	1	X51732	DNA encoding a hum
625	45	2.6	986	1	X51735	DNA encoding a hum
626	45	2.6	3495	1	X33814	Coding sequence fo
627	45	2.6	1493	1	X33994	Human osteopontin

628	45	2.6	1547	1	X33448	Oryza sativa L. pi
629	45	2.6	1984	1	X52252	Protein PRO271 CDN
630	45	2.6	1503	1	X52253	Protein PRO272 CDN
631	45	2.6	2822	1	X52222	Protein PRO228 CDN
632	45	2.6	1060	1	X37464	Human secreted pro
633	45	2.6	1448	1	X37462	Human secreted pro

ALIGNMENTS

RESULT	1	
ID	080521	
AC	080521: standard; cDNA; 1737 BP.	
DT	18-JUL-1995 (first entry)	
DE	Human monocyte PFA4R CDNA.	
KW	Interleukin-8 receptor; IL-8 receptor; PFA4R;	
KW	platelet factor superfamily receptor; monocyte; chemotactic;	
KW	inflammation; inflammatory disease; arthritis; emphysema; cystic;	
KW	fibrosis; colitis; bronchitis; meningitis; therapeutic; ss.	
OS	Homo sapiens.	
FT	Key	Location/Qualifiers
FT	cds	91..1149
FT		/*tag= a
PN	W09428931-A.	
PD	22-DEC-1994.	
PF	07-JUN-1994.	
PR	11-JUN-1993; US-076093.	
PA	(GETH) GENENTECH INC.	
PI	Chuntharapai A, Hebert C, Kim KJ, Lee J;	
DR	WPI: 95-036114/05.	
DR	P-PSDB: R68812.	
PT	Treatment of inflammatory disorders - by administering an	
PT	antibody capable of binding a platelet factor 4 superfamily	
PT	receptor polypeptide	
PS	Disclosure; Page 54-56; 83pp; English.	
CC	2 PFA4R members were identified by probing lambda libraries from	
CC	human monocyte-like cell line HL-60 and human peripheral blood	
CC	lymphocytes using a large fragment of IL-8 receptor DNA (full	
CC	sequence given in Q80520). The nucleotide sequences of the 2	
CC	PFA4Rs are given in Q80521 and Q80522, and their respective	
CC	amino acid sequences in R68812 and R68813.	
SQ	Sequence 1737 BP; 454 A; 411 C; 373 G; 499 T;	

Query Match	100.0%;	Score 1737;	DB 1;	Length 1737;
Best Local Similarity	100.0%;	Pred. NO. 0;		
Matches 1737;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	GAATTCAGTGTGCTGCGCGCGCGCGCAAGTGACGCGCGAGGCGCTGAGTCTCAGTA	60
DB	1	CAATTCAGTGTGCTGCGCGCGCGCGCGCAAGTGACGCGCGAGGCGCTGAGTCTCAGTA	60
QY	61	GCCACCGCATCTGGAAGAACCGGTTACCATGGAGGGATCAGTATATACACTTCAGAT	120
DB	61	GCCACCGCATCTGGAAGAACCGGTTACCATGGAGGGATCAGTATATACACTTCAGAT	120
QY	121	AACATACCGGAAATGGGCTCAGGAGCATGACATCCATGAAGAACCTGTTCCGT	180
DB	121	AACATACCGGAAATGGGCTCAGGAGCATGACATCCATGAAGAACCTGTTCCGT	180
QY	181	GAAGAAATGCTAATTTCAATTAATTTCTTCTGCGCCACCATCTACTCCATCATCTTCTTA	240
DB	181	GAAGAAATGCTAATTTCAATTAATTTCTTCTGCGCCACCATCTACTCCATCATCTTCTTA	240
QY	241	ACTGCGATTGTGGGAATGATGGTTCATCTCGGTCATGGGTACTACAGAGAAGCTGAGA	300
DB	241	ACTGCGATTGTGGGCAATGGAATGGTTCATCTCTGTCATAGGGTTACCAAGAAACTGAGA	300
QY	301	AGCATGACGACAAAGTACAGGCTGCACCTGTAGTGGCGGACCTCTCTTGTCATCAG	360
DB	301	AGCATGACGACAAAGTACAGGCTGCACCTGTAGTGGCGGACCTCTCTTGTCATCAG	360

```

QY 361 CTTCCTCTGGGCACTGATGCGCGTGGCAAACTGCTACTTTGGGAACCTCTATGCAAG 420
DB 361 CTTCCTCTGGGCACTGATGCGCGTGGCAAACTGCTACTTTGGGAACCTCTATGCAAG 420
QY 421 GCAGTCATCTCATCTACAGTCACAGTCACAGTCACAGTCACAGTCACAGTCACAGTCAC 480
DB 421 GCAGTCATCTCATCTACAGTCACAGTCACAGTCACAGTCACAGTCACAGTCACAGTCAC 480
QY 481 AGAGTCGAGCGCTACCTGCGCATCTGTCAGCGCACCAACAGTCAGAGGCCAAGAGCTG 540
DB 481 AGAGTCGAGCGCTACCTGCGCATCTGTCAGCGCACCAACAGTCAGAGGCCAAGAGCTG 540
QY 541 TTGGCTGAAAAAGTGCTATGTTGGGCTGGAATCCCTGCGCTCTGCTACATATCC 600
DB 541 TTGGCTGAAAAAGTGCTATGTTGGGCTGGAATCCCTGCGCTCTGCTACATATCC 600
QY 601 GACTTCATCTTTGCCAAGCTAGTGAAGGAGATGAGATATATCTGAGCCGCTTAC 660
DB 601 GACTTCATCTTTGCCAAGCTAGTGAAGGAGATGAGATATATCTGAGCCGCTTAC 660
QY 661 CCCATGACTTGTGGTGTGTGTGTTCCAGTTTCAGCAGCATATGCTGCGCTTATCCTG 720
DB 661 CCCATGACTTGTGGTGTGTGTGTTCCAGTTTCAGCAGCATATGCTGCGCTTATCCTG 720
QY 721 CCTGGTATTCATCTCTCTCTCTATTCATATCATCTCCAGCTGTCACAGCTCCAG 780
DB 721 CCTGGTATTCATCTCTCTCTCTATTCATATCATCTCCAGCTGTCACAGCTCCAG 780
QY 781 GGGCACCAGAGCGCAAGCGCCCTCAAGACACAGTCATCTCATCTGCTTCTTTCGCC 840
DB 781 GGGCACCAGAGCGCAAGCGCCCTCAAGACACAGTCATCTCATCTGCTTCTTTCGCC 840
QY 841 TGTGGCTGCTTACTATCATTTGGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 841 TGTGGCTGCTTACTATCATTTGGATGATGATGATGATGATGATGATGATGATGATGATG 900
QY 901 AAGCAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
DB 901 AAGCAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 961 GCTTCTCTCCACTGTTGTCTGTAACCCCACTCTATGCTTCTTGGAGCCAAATTTAA 1020
DB 961 GCTTCTCTCCACTGTTGTCTGTAACCCCACTCTATGCTTCTTGGAGCCAAATTTAA 1020
QY 1021 ACCTTGTCCAGCAGCAGCTACCTCTGTGAGCAGAGGCTCCAGCTCAAGATTCCTCC 1080
DB 1021 ACCTTGTCCAGCAGCAGCTACCTCTGTGAGCAGAGGCTCCAGCTCAAGATTCCTCC 1080
QY 1081 AAAGGAAAGCGAGTGGAGCATTCATCTGTTCCAGTGTGAGTCTCAAGTTTTCAC 1140
DB 1081 AAAGGAAAGCGAGTGGAGCATTCATCTGTTCCAGTGTGAGTCTCAAGTTTTCAC 1140
QY 1141 TCCAGCTACACAGATGTAAGAAAGCTTTTTTATACGATTAATTAATTTTAAAGTT 1200
DB 1141 TCCAGCTACACAGATGTAAGAAAGCTTTTTTATACGATTAATTAATTTTAAAGTT 1200
QY 1201 ACACATTTTTCAGATATTAAGAAAGCTCAACAATTTTACAGTTTATGCTGTGAT 1260
DB 1201 ACACATTTTTCAGATATTAAGAAAGCTCAACAATTTTATGCTGTGAT 1260
QY 1261 TTTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
DB 1261 TTTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
QY 1321 TTTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
DB 1321 TTTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
QY 1381 GTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
DB 1381 GTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440

```

```

QY 1441 TAAAGCTAGAAATGATCCCGAGCTGTTATGATAGATTAATCTCTCCATCCCGTGAAC 1500
DB 1441 TAAAGCTAGAAATGATCCCGAGCTGTTATGATAGATTAATCTCTCCATCCCGTGAAC 1500
QY 1501 GTTTTCTCTGTTCTTAAAGCTGATTTTGTGTAGAAAGATGAGCAGTATTAACCAAGCCC 1560
DB 1501 GTTTTCTCTGTTCTTAAAGCTGATTTTGTGTAGAAAGATGAGCAGTATTAACCAAGCCC 1560
QY 1561 AAAGTGTATGAAGATCTGTTTTCAGTTTCAGAGTGGGTTGATTTTCAGCACCTAC 1620
DB 1561 AAAGTGTATGAAGATCTGTTTTCAGTTTCAGAGTGGGTTGATTTTCAGCACCTAC 1620
QY 1621 AGTGTACAGCTTGTATTAAGTTGTTAATAAAGTACATGTTAACTTAAAAA 1680
DB 1621 AGTGTACAGCTTGTATTAAGTTGTTAATAAAGTACATGTTAATAA 1680
QY 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACACAGCTGAATTC 1737
DB 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCCGACACAGCTGAATTC 1737

RESULT 2
ID 099007 standard; cDNA; 1737 BP.
AC 099007:
DT 26-MAR-1996 (first entry)
DE Chemokine superfamily receptor coding sequence.
KW Interleukin: IL-8; inflammation; psoriasis; dermatitis;
KW Rheumatoid arthritis; inflammatory bowel disease;
KW Chronic lung inflammation; treatment; antibody;
KW affinity purification; detection; ss.
OS Homo sapiens.
PN US5440021-A.
PD 08-AUG-1995.
PR 29-MAR-1991: 677211.
PR 25-FEB-1994: US-677211.
PR (CHUN)/ CHUNTHARAPAI A.
PA (HEBE)/ HEBERT C.
PA (KIMK)/ KIM K J.
PA (LEEJ)/ LEE J.
PI Chuntharapai A., Hebert C., Kim KJ, Lee J;
PI WPI: 95-283151/37.
DR P-PSDB: R80757.
PT New antibodies against interleukin 8 type B receptor - used to treat
PT or prevent inflammation, also for detecting receptor expression and
PT purification.
PS Example 2: Columns 47-50: 62bp; English.
CC Antibodies directed against the interleukin-8 receptor B can be used
CC to treat or prevent inflammation e.g. psoriasis, dermatitis.
CC Rheumatoid arthritis and particularly inflammatory bowel disease and
CC chronic lung inflammation. When immobilised, these antibodies may
CC be used to detect interleukin-8 receptor B expression in cells and
CC tissues and for affinity purification of interleukin-8 receptor B
CC from cells. This sequence is an additional chemokine superfamily
CC receptor which was identified by probing lambda libraries of genomic
CC DNA from a human monocyte-like cell line (L-60) and human peripheral
CC blood lymphocytes using a large fragment of the interleukin-8 type
CC A receptor DNA (See Q99006).
SO Sequence 1737 BP; 454 A; 411 C; 373 G; 499 T;

Query Match 100.0%; Score 1737; DB 1: Length 1737;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1737; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAATTCAGTGTGCTGGGCGGCGGCAAGTGAAGCCGAGGCTGAGTCCAGTA 60
DB 1 GAATTCAGTGTGCTGGGCGGCGGCAAGTGAAGCCGAGGCTGAGTCCAGTA 60
QY 61 GCCACCGCATCTGAGAACCGAGCGTTACCATGAGAGGATCAGTATPACACTTCAAT 120
DB 61 GCCACCGCATCTGAGAACCGAGCGTTACCATGAGAGGATCAGTATPACACTTCAAT 120

```

OY	121	AACATACCCAGGAAATGGGCTAGGGGACTATGATCTCATAGAGAACCTCTTGCCG	180
Dp	121	AACATACCCAGGAAATGGGCTAGGGGACTATGATCTCATAGAGAACCTCTTGCCG	180
OY	181	GAGAGAAATGCTAATTTCATTAATAATCTTCCTGCCACCACATGCTACAGCATCACTCTCTA	240
Dp	181	GAGAGAAATGCTAATTTCATTAATAATCTTCCTGCCACCACATGCTACAGCATCACTCTCTA	240
OY	241	ACTGGCAATTGTGGGCAATGGATTGGTCAATCTGGTCAATGGGTTACAGAGAAACTGAGA	300
Dp	241	ACTGGCAATTGTGGGCAATGGATTGGTCAATCTGGTCAATGGGTTACAGAGAAACTGAGA	300
OY	301	AGCATGACGGGCAAGTACAGGCTGACACCTGTACGTGAGCCGACCTGCTCTTGTACATCAG	360
Dp	301	AGCATGACGGGCAAGTACAGGCTGACACCTGTACGTGAGCCGACCTGCTCTTGTACATCAG	360
OY	361	CTTCCCTCTTGCGGAGTTGATGCCGTGGCAACTGGTACTTTGGGAACHTTCTATGACAG	420
Dp	361	CTTCCCTCTTGCGGAGTTGATGCCGTGGCAACTGGTACTTTGGGAACHTTCTATGACAG	420
OY	421	GCAATCCATGTCATCTACACAGTCAACTTACAGCAGTGTCTCATCTGGCCTTATC	480
Dp	421	GCAATCCATGTCATCTACACAGTCAACTTACAGCAGTGTCTCATCTGGCCTTATC	480
OY	481	AGTGTGACCCGCTACCTGGCCATCTGTCCAGCCACCAAGTACAGAGGCAAGAACTG	540
Dp	481	AGTGTGACCCGCTACCTGGCCATCTGTCCAGCCACCAAGTACAGAGGCAAGAACTG	540
OY	541	TTGGCTGAAAAGTGGTCTATGTTGGGCTGTGATCCTCTCCCTCTGCTACTATTC	600
Dp	541	TTGGCTGAAAAGTGGTCTATGTTGGGCTGTGATCCTCTCCCTCTGCTACTATTC	600
OY	601	GACTTCATCTTTGGCAACGTCAGTGAAGGCAATGACAGATATATCTGTGACCGCTTAC	660
Dp	601	GACTTCATCTTTGGCAACGTCAGTGAAGGCAATGACAGATATATCTGTGACCGCTTAC	660
OY	661	CCCAATGACTTGTGGGTGTTGTTGCCATTTCAGACACATCATGGTTGGCCTTATTC	720
Dp	661	CCCAATGACTTGTGGGTGTTGTTGCCATTTCAGACACATCATGGTTGGCCTTATTC	720
OY	721	CGTGTATGTGATACCTCTGCTCTGCTATGTGATATATATCTCCAAGCTGTGCACATC	780
Dp	721	CGTGTATGTGATACCTCTGCTCTGCTATGTGATATATATCTCCAAGCTGTGCACATC	780
OY	781	GGCACCCAGAGGCGCAGAGGCCCTCAGACACAGTATCTCATCTGGCTTTCTGCC	840
Dp	781	GGCACCCAGAGGCGCAGAGGCCCTCAGACACAGTATCTCATCTGGCTTTCTGCC	840
OY	841	TGTTGGCTGCCCTTACTACATTTGGGATCAGCATGACTCCTTATCTCTCTGGAAATC	900
Dp	841	TGTTGGCTGCCCTTACTACATTTGGGATCAGCATGACTCCTTATCTCTCTGGAAATC	900
OY	901	AAGCAAGGGTGTAGTTGAGAACACTGTGCAACAAATGATTTTCATCAGCAGGCCCTA	960
Dp	901	AAGCAAGGGTGTAGTTGAGAACACTGTGCAACAAATGATTTTCATCAGCAGGCCCTA	960
OY	961	GCCTTCTTCCACTGTTGTGTGAACCCCATCTCTATGCTTTCTCTGAGCCAAATTTAA	1020
Dp	961	GCCTTCTTCCACTGTTGTGTGAACCCCATCTCTATGCTTTCTCTGAGCCAAATTTAA	1020
OY	1021	ACCTCTGCCAGACGCACTCACTCTGTGAGAGAGGGTCCAGGCTCAAGTCTCTCC	1080
Dp	1021	ACCTCTGCCAGACGCACTCACTCTGTGAGAGAGGGTCCAGGCTCAAGTCTCTCC	1080
OY	1081	AAAGCAAGCGAGGTGACATTCATCTGTTTCCAGTGAATCTGAGATCTTCAAGTTTAC	1140
Dp	1081	AAAGCAAGCGAGGTGACATTCATCTGTTTCCAGTGAATCTGAGATCTTCAAGTTTAC	1140
OY	1141	TCCAGCTAACACAGATGTAAGACATTTTTTTTATACGATTAATAACTTTTTTAAGT	1200
Dp	1141	TCCAGCTAACACAGATGTAAGACATTTTTTTTATACGATTAATAACTTTTTTTAAGT	1200

Oy	1201	ACACATTTTTCAGATATAAAAGCATGACCATTGTGTACAGTTTTATTGGCTGTGGAT	1260
Dd	1201	ACAACATTTTTCAGATATAAAAGCATGACCATTGTGTACAGTTTTATTGGCTGTGGAT	1260
Oy	1261	TTTGCTCGTGGTCTCTTAAGTTTGTGCAGATTAAATGACTTAATATAAATTTT	1320
Dd	1261	TTTGCTCGTGGTCTCTTAAGTTTGTGCAGATTAAATGACTTAATATAAATTTT	1320
Oy	1321	TTTGTTCAATATGATGTGTCTTAGGCAGAGACTGTGGCCAGTTCTTAAATGCTGAT	1380
Dd	1321	TTTGTTCAATATGATGTGTCTTAGGCAGAGACTGTGGCCAGTTCTTAAATGCTGAT	1380
Oy	1381	GTCGCGGTGTGGACCTGTAAAAAGGAAGCAGTGAACATTCAGAAGGCTGTAGGAATCACG	1440
Dd	1381	GTCGCGGTGTGGACCTGTAAAAAGGAAGCAGTGAACATTCAGAAGGCTGTAGGAATCACG	1440
Oy	1441	TAAAGCTAGAAATATGCCCGAGCTGTTTATGCATAGATATCTCTCATTCCTCCGTGAAAC	1500
Dd	1441	TAAAGCTAGAAATATGCCCGAGCTGTTTATGCATAGATATCTCTCATTCCTCCGTGAAAC	1500
Oy	1501	GTTTTTCTGTCTTTAAGACGTATTTTGGCTGAGAAATGGCACTTATAACCAAAGCCC	1560
Dd	1501	GTTTTTCTGTCTTTAAGACGTATTTTGGCTGAGAAATGGCACTTATAACCAAAGCCC	1560
Oy	1561	AAAGTGATAGAAATGCTGTTTTCAGTTTTCAGAGATGGGTTGATTTACGACCTAC	1620
Dd	1561	AAAGTGATAGAAATGCTGTTTTCAGTTTTCAGAGATGGGTTGATTTACGACCTAC	1620
Oy	1621	AGTGTACAGTCTGTATTAAAGTGTTAATAAAGTACATGTTAAACTTAAAAA	1680
Dd	1621	AGTGTACAGTCTGTATTAAAGTGTTAATAAAGTACATGTTAAACTTAAAAA	1680
Oy	1681	AAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCGCGCAGACACTGGAATTC	1737
Dd	1681	AAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCGCGCAGACACTGGAATTC	1737
<hr/>			
RESULT	3		
ID	Q29506	Q29506 standard; DNA; 1737 BP.	
AC	Q29506:		
DT	12-MAR-1993	(first entry)	
DE	New platelet factor 4 receptor superfamily member PF4ARI.		
KW	IL-8R; G-protein coupled receptor family; rhodopsin superfamily;		
KM	pro-inflammatory cytokine; 8rr.20.15; ss.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	91..1149	
FT		/tag= a	
FT		/product= PF4ARI	
PN	M09217497-A.		
PD	15-OCT-1992.		
PF	23-MAR-1992:	U02317.	
PPR	29-MAR-1991:	US-677211.	
PPR	19-DEC-1991:	US-810782.	
PA	(GENE) GENENTECH INC.		
P1	Holmes WE, Lee J, Wood WT;		
DR	WPJ: 92-366191/44.		
DR	P-PDB: R27792.		
PT	Isolated human platelet factor 4 super-family receptor		
PT	polypeptide and corresp. antibodies and DNA - useful as		
PT	diagnostic and screening agents, and for treating inflammation or		
PT	PF4AR-mediated disorders		
PS	Claim 7; Fig 4; 78pp; English.		
CC	The IL-8 receptor cDNA sequence was isolated (see Q29505) and a		
CC	874bp sub-fragment of the coding sequence was used as a probe to		
CC	screen human cell line Hf60 and human peripheral blood lymphocyte		
CC	cDNA libraries. Two new gene sequences were found that are clearly		
CC	related to the IL-8 receptor. One of these was contained in		
CC	combined clone 8rr.20.15 and is predicted to encode an amino acid		
CC	sequence which is 34% identical with both the high and low affinity		
CC	IL-8 receptors. See also Q37107.		
SO	Sequence	1737 BP: 457 A: 412 C: 370 G: 498 T:	

Query Match 88.3%; Score 1533; DB 1; Length 1737;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1733; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAATTCAGTGTGCTGGCGGCGGCGGCAAAAGTACGCCGAGGAGCCGAGTGTCTCCAGTA 60
DB 1 GAATTCAGTGTGCTGGCGGCGGCGGCAAAAGTACGCCGAGGAGCCGAGTGTCTCCAGTA 60
QY 61 GCCACCGCATCTGAGAGAACGCGGTTACCATGAGGGGATCATATATACACTTCAGAT 120
DB 61 GCCACCGCATCTGAGAGAACGCGGTTACCATGAGGGGATCATATATACACTTCAGAT 120
QY 121 AACTACACCGAAGAAATGGGCTCAGGGGACTATGACTCCATGAAGAAACCTGTTCCGT 180
DB 121 AACTACACCGAAGAAATGGGCTCAGGGGACTATGACTCCATGAAGAAACCTGTTCCGT 180
QY 131 GAGAAATATGCTAATTTCAATATAAATCTTCTGCGCCACATCTACTCATCATCTTTCTTA 240
DB 131 GAGAAATATGCTAATTTCAATATAAATCTTCTGCGCCACATCTACTCATCATCTTTCTTA 240
QY 241 ACTGCAATGTGGGCAATGGATTGTCATCCTGCTCATGGGTTACGAGAAACTGAGA 300
DB 241 ACTGCAATGTGGGCAATGGATTGTCATCCTGCTCATGGGTTACGAGAAACTGAGA 300
QY 301 AGCATGACGACAGTACAGGCTGACCTGTCAGTGGCGGACCTCTTTGTCATCAG 360
DB 301 AGCATGACGACAGTACAGGCTGACCTGTCAGTGGCGGACCTCTTTGTCATCAG 360
QY 361 CTTCCTCTTGGGAGTGTAGTCCGCGCAAACTGTACTTTGGGAACTTCCATGCAAG 420
DB 361 CTTCCTCTTGGGAGTGTAGTCCGCGCAAACTGTACTTTGGGAACTTCCATGCAAG 420
QY 421 GCAGTCCATGTCATCTACACAGTCAACCTCTACAGCACTGTCCTCATCTGGCTTCATC 480
DB 421 GCAGTCCATGTCATCTACACAGTCAACCTCTACAGCACTGTCCTCATCTGGCTTCATC 480
QY 481 AGTGTGACCCGCTACCTGGCCATGTCACGCGCACACAGTCAAGAGGAGCTG 540
DB 481 AGTGTGACCCGCTACCTGGCCATGTCACGCGCACACAGTCAAGAGGAGCTG 540
QY 541 TTGCTGAAAAAGTGTCTATGTTGGCTGTGATCCCTGCGCTCTGCTGACTATTCCTC 600
DB 541 TTGCTGAAAAAGTGTCTATGTTGGCTGTGATCCCTGCGCTCTGCTGACTATTCCTC 600
QY 601 GACTTCATCTTTGCCAAGCTCAGTACGAGTACAGATATATCTGTGACCCGCTTCTAC 660
DB 601 GACTTCATCTTTGCCAAGCTCAGTACGAGTACAGATATATCTGTGACCCGCTTCTAC 660
QY 661 CCCAATGCTGTGGGAGGTTGTGTTCCAGTTTCAGACATCATGGTGGCTTATCCTG 720
DB 661 CCCAATGCTGTGGGAGGTTGTGTTCCAGTTTCAGACATCATGGTGGCTTATCCTG 720
QY 721 CCTGATATGTCATCTCTCTCTGCTATTTGATATCATCTCCAAAGCTGTACACTCCAAAG 780
DB 721 CCTGATATGTCATCTCTCTCTGCTATTTGATATCATCTCCAAAGCTGTACACTCCAAAG 780
QY 781 GGGCACCAGAAAGCGCAAGGCGCTCAAGACACAGTACATCTCTGCTTTCTTGCGC 840
DB 781 GGGCACCAGAAAGCGCAAGGCGCTCAAGACACAGTACATCTCTGCTTTCTTGCGC 840
QY 841 TGTGGCTGCGCTTACTACATTTGGGATCGAGATCGACTCTCTTATCTCTCTGGAATCATC 900
DB 841 TGTGGCTGCGCTTACTACATTTGGGATCGAGATCGACTCTCTTATCTCTCTGGAATCATC 900
QY 901 AAGCAAGGATGTAGTTGAGAACACTGTGCAAGATGATTTCCATACCGAGGCCCTA 960
DB 901 AAGCAAGGATGTAGTTGAGAACACTGTGCAAGATGATTTCCATACCGAGGCCCTA 960
QY 961 GCTTCTTCCACTGTGTGTGAACCCATCTCTATGCTTTCTTGAGACCAATTTTAA 1020
DB 961 GCTTCTTCCACTGTGTGTGTGAACCCATCTCTATGCTTTCTTGAGACCAATTTTAA 1020

QY 1021 ACTCTGCCAGACGACACTCACTCTGTGACAGAGGGTCCAGCTCAAGATCCTCTCC 1080
DB 1021 ACTCTGCCAGACGACACTCACTCTGTGACAGAGGGTCCAGCTCAAGATCCTCTCC 1080
QY 1081 AAGGAAAGCGAGGTGGACATTCATCTGTTTCCATCGATCTGACCTTCAAGTTTCCAC 1140
DB 1081 AAGGAAAGCGAGGTGGACATTCATCTGTTTCCATCGATCTGAGCTTCAAGTTTCCAC 1140
QY 1141 TCCAGTAAACAGATGTAAAGACTTTTATACGATTAATACTTTTAAAGTT 1200
DB 1141 TCCAGTAAACAGATGTAAAGACTTTTATACGATTAATACTTTTAAAGTT 1200
QY 1201 ACACATTTTTCAGATATAAAGACTACCAATATTTGATACGTTTATTTGCTTGGAT 1260
DB 1201 ACACATTTTTCAGATATAAAGACTACCAATATTTGATACGTTTATTTGCTTGGAT 1260
QY 1261 TTTTGTCTGTGTTTCTTTAGTTTGTGAGTTTAAATGACTTATTAATAATTTT 1320
DB 1261 TTTTGTCTGTGTTTCTTTAGTTTGTGAGTTTAAATGACTTATTAATAATTTT 1320
QY 1321 TTTGTTTCATATTGATGTGTCTAGCAGAGACCTGTGSCCAAGTTCTTAGTCTGTAT 1380
DB 1321 TTTGTTTCATATTGATGTGTCTAGCAGAGACCTGTGSCCAAGTTCTTAGTCTGTAT 1380
QY 1381 GTCTCGTGTAGGACTGTAGAAAGGAGACTGAACATTTCCAGAGCGTGTGATTCAG 1440
DB 1381 GTCTCGTGTAGGACTGTAGAAAGGAGACTGAACATTTCCAGAGCGTGTGATTCAG 1440
QY 1441 TAAAGTAGAAATGATCCCGAGCTTTATGCAATAGTAAATCTCTCATTCCTGGTGAAC 1500
DB 1441 TAAAGTAGAAATGATCCCGAGCTTTATGCAATAGTAAATCTCTCATTCCTGGTGAAC 1500
QY 1501 GTTTTCTCTGTTCTTAAAGCTGATTTTGTGTGAGAAAGTGGCACTTATAACCAAGCCC 1560
DB 1501 GTTTTCTCTGTTCTTAAAGCTGATTTTGTGTGAGAAAGTGGCACTTATAACCAAGCCC 1560
QY 1561 AAAGTGTATGAAATCTGTTTTCAGTTTCAGAGTGGGTTGATTTTCAGACCTTAC 1620
DB 1561 AAAGTGTATGAAATCTGTTTTCAGTTTCAGAGTGGGTTGATTTTCAGACCTTAC 1620
QY 1621 AGTGTACAGTGTGTATTAAGTTGTATATAAAGTACATGTATACTTAAATAAAAAA 1680
DB 1621 AGTGTACAGTGTGTATTAAGTTGTATATAAAGTACATGTATACTTAAATAAAAAA 1680
QY 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGGCCCGCCAGCACACTGGAATTC 1737
DB 1681 AAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGGCCCGCCAGCACACTGGAATTC 1737

RESULT 4
X15882
ID X15882 standard; cDNA: 1944 BP.
AC X15882:
DE 12-MAY-1999 (first entry)
DE cDNA encoding G-protein coupled receptor polypeptide designated CXCR4B.
KW G-protein coupled receptor; CXCR4B; human; splice variant;
KW chemokine receptor; CXCR4; vaccine; gene therapy; HIV-1; HIV-1;
KW cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;
KW Parkinson's disease; acute heart failure; hypotension; hypertension;
KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
KW stroke; ulcer; allergy; benign prostatic hyperplasia; migraine;
KW vomiting; psychotic disorder; neurological disorder; anxiety;
KW schizophrenia; manic depression; delirium; dementia; mental retardation;
KW dyskinesias; Huntingtons disease; Gilles de la Tourette syndrome; ss.
OS Homo sapiens.
FH Key location/Qualifiers
FT CDS 336..1406
FT FT /*tag= a
FT FT misc_feature 361
FT FT /*tag= b
FT FT /note= "splice acceptor site"
PN EP-89780-A2.

24-FEB-1999.
PF 07-AUG-1998: 306324.
PR 24-JUL-1998: US-056601.
PR 20-AUG-1997: US-056601.
PA (SMIK) SMITHKLINE BEECHAM CORP.
PI Gupta SK, Pillarisetla K;
DR WPI: 99-134643/12.
DR P-PSDB: W97362.
PT New G protein coupled receptor (CXCR4B) polypeptide and
PI polynucleotide, human splice variant of a chemokine receptor -
PT useful as diagnostic reagents and for prevention and treatment of
PT HIV infection, cancer, stroke and dementia
PS Claim 2: Page 16-17: 24pb: English.
CC The present sequence encodes a G-protein coupled receptor polypeptide
CC designated CXCR4B. CXCR4B polynucleotides and polypeptides are useful for
CC diagnosing susceptibility to diseases by detecting mutations or
CC polymorphisms in the CXCR4B gene or analysing for the presence or amount
CC of CXCR4B polypeptide expressed in a patient. CXCR4B polypeptides and
CC polynucleotides are also useful for screening for antagonists and
CC agonists which can be used to treat conditions associated with CXCR4B
CC polypeptide imbalance. CXCR4B polypeptides can be administered directly
CC (as a vaccine) or via a vector (gene therapy) to prevent disease.
CC Diseases diagnosed, prevented and treated include: HIV-1 and HIV-1
CC infections; cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;
CC Parkinson's disease; acute heart failure; hypotension; hypertension;
CC urinary retention; osteoporosis; angina pectoris; myocardial infarction;
CC stroke; ulcers; allergies; benign prostatic hypertrophy; migraine;
CC vomiting; psychotic and neurological disorders, including anxiety,
CC schizophrenia, manic depression, delirium, dementia and severe mental
CC retardation; and dyskinesias, such as Huntington's disease or Gilles de
CC la Tourette syndrome. CXCR4B polypeptides are useful for mapping genes
CC to chromosomes, allowing gene inheritance to be studied through linkage
CC analysis.
SQ Sequence 1944 BP: 497 A: 437 C: 412 G: 396 T:

Query Match 87.7%; Score 1523; DB 1; Length 1944;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1523; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ATATACCTTCACTACACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG 165
DB 363 ATATACCTTCACTACACTACACCGAGGAAATGGGCTCAGGGGACTATGACTCCATGAAG 422
QY 166 GAACCCGTTCCGTGAGAAATGCTAATTTCAATAAATCTTCCGCCACCATATAC 225
DB 423 GAACCCGTTCCGTGAGAAATGCTAATTTCAATAAATCTTCCGCCACCATATAC 482
QY 226 TCATCATCTTCTTAAGTGGCATGTGGCAATGATGGTCAATCCTGGTCAATGGATTAC 285
DB 483 TCATCATCTTCTTAAGTGGCATGTGGCAATGATGGTCAATCCTGGTCAATGGATTAC 542
QY 285 CAAAGAACTGGAAGCATGACGAGCAATACAGGCTGCACCTGTGAGTGGGACCTC 345
DB 543 CAAAGAACTGGAAGCATGACGAGCAATACAGGCTGCACCTGTGAGTGGGACCTC 602
QY 346 CTCTTTGTCAACGCTTCCCTTCTGGGCAAGTGGTCCCTGGCAACTGGTACTTTGGG 405
DB 603 CTCTTTGTCAACGCTTCCCTTCTGGGCAAGTGGTCCCTGGCAACTGGTACTTTGGG 662
QY 406 AACTTCTATGCAAGGAGTGCATGTCAATACAGACTCAACCTCTACAGAGTGTCTC 465
DB 663 AACTTCTATGCAAGGAGTGCATGTCAATACAGACTCAACCTCTACAGAGTGTCTC 722
QY 466 ATCTGCGCTTCACTACGCTGAGCCGCTACCTGGCCATCTCTCAACGCGACACAGTCAAG 525
DB 723 ATCTGCGCTTCACTACGCTGAGCCGCTACCTGGCCATCTCTCAACGCGACACAGTCAAG 782
QY 526 AGCCCAAGAGAGCTGTTGGCTGAAAAGTGGTATATGTTGGCGTCTGGATCCCTGCCCTC 585
DB 783 AGCCCAAGAGAGCTGTTGGCTGAAAAGTGGTATATGTTGGCGTCTGGATCCCTGCCCTC 842

QY 586 CTGCTGACTATTCGCCGACTTCACTTTCGCCAAGCTGACGAGCATGACATATATC 645
DB 843 CTGCTGACTATTCGCCGACTTCACTTTCGCCAAGCTGACGAGCATGACATATATC 902
QY 646 TGTGACCGCTTCAACCCCAATGACTTGGGTGGTGTGTTCAGTTTCAGACATCATG 705
DB 903 TGTGACCGCTTCAACCCCAATGACTTGGGTGGTGTGTTCAGTTTCAGACATCATG 962
QY 706 GTTGGCTTATCTGCTGCTGATTTGATCATCTGTCTGTCTATTCATATCATCTCAAG 765
DB 963 GTTGGCTTATCTGCTGCTGATTTGATCATCTGTCTGTCTATTCATATCATCTCAAG 1022
QY 766 CTGTCACATCCCAAGGGCCACGCAAGCGAAGGCCCTCAAGACACAGATCATCTCATC 825
DB 1023 CTGTCACATCCCAAGGGCCACGCAAGCGAAGGCCCTCAAGACACAGATCATCTCATC 1082
QY 826 CTGGCTTTCTCGGCTTGTGGTGGCTTACTACATTTGGGATCAGCATGACTCCTTCATC 885
DB 1083 CTGGCTTTCTCGGCTTGTGGTGGCTTACTACATTTGGGATCAGCATGACTCCTTCATC 1142
QY 886 CTCTGGAATTCATCAAGCAAGGTGTGAGTTTGAGACACTGTGCACAGTGGATTCC 945
DB 1143 CTCTGGAATTCATCAAGCAAGGTGTGAGTTTGAGACACTGTGCACAGTGGATTCC 1202
QY 946 ATACCGAGGCGCCAGCTTCTTCCACTGTGTCTGACACCCCATCTCTATGCTTCTCT 1005
DB 1203 ATACCGAGGCGCCAGCTTCTTCCACTGTGTCTGACACCCCATCTCTATGCTTCTCT 1262
QY 1006 GGAAGCCAAATTTAAACCTGTGCCACACGACACTCCTGTGTGACAGAGGCTCCAGC 1065
DB 1263 GGAAGCCAAATTTAAACCTGTGCCACACGACACTCCTGTGTGACAGAGGCTCCAGC 1322
QY 1066 CTCAAGATCTCTCCAAAGGAAAGCGAGGTGGACATTCACTGTTCACATGACTGAG 1125
DB 1323 CTCAAGATCTCTCCAAAGGAAAGCGAGGTGGACATTCACTGTTCACATGACTGAG 1382
QY 1126 TCTTCAAGTTTCACTCCACCTCAACACAGATGTAAGACTTTTATACGATAATA 1185
DB 1383 TCTTCAAGTTTCACTCCACCTCAACACAGATGTAAGACTTTTATACGATAATA 1442
QY 1186 ACTTTTATTAAGTTACATATTTTCAGATATAAAGACTGACCAATATTTGACAGTTT 1245
DB 1443 ACTTTTATTAAGTTACATATTTTCAGATATAAAGACTGACCAATATTTGACAGTTT 1502
QY 1246 TATTGCTGTGGATTTTGTCTGTGTCTTCTTGTGAGTTTGTGAGTTAATTGACTTA 1305
DB 1503 TATTGCTGTGGATTTTGTCTGTGTCTTCTTGTGAGTTTGTGAGTTAATTGACTTA 1562
QY 1306 TTTATATAATTTTATTTTGTTCATATATGATGTGTCTGAGCAGGACCTGTGGCAAGT 1365
DB 1563 TTTATATAATTTTATTTTGTTCATATATGATGTGTCTGAGCAGGACCTGTGGCAAGT 1622
QY 1366 TCTTTAGTGTGATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1425
DB 1623 TCTTTAGTGTGATGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1682
QY 1426 GTGAGTGAATCAAGTAAAGCTGAAATGATCCCGAGCTGTTATGATGATGATATCTCT 1485
DB 1583 GTGAGTGAATCAAGTAAAGCTGAAATGATCCCGAGCTGTTATGATGATGATATCTCT 1742
QY 1486 CCAATCCCGTGAAGCTTTTCTGTCTTAAGACGTGATTTTGTGTAAGAGATGGCAC 1545
DB 1743 CCAATCCCGTGAAGCTTTTCTGTCTTAAGACGTGATTTTGTGTAAGAGATGGCAC 1802
QY 1546 TTTATACCAAAAGCCCAAAGTGTATGAAATGCTGTGTGTGTGTGTGTGTGTGTGTGT 1605
DB 1803 TTTATACCAAAAGCCCAAAGTGTATGAAATGCTGTGTGTGTGTGTGTGTGTGTGTGT 1862
QY 1606 GATTTCAGACCTTACAGTGTACA 1628
DB 1863 GATTTCAGACCTTACAGTGTACA 1885

RESULT 5
ID V18357 standard: DNA; 1317 BP.
AC V18357:
DT 25-SEP-1998 (first entry)
DE Human RM3 seven transmembrane (7TM) receptor cDNA.
KW V8: Placenta; seven transmembrane receptor; 7TM; signal transduction;
KM immunology; inflammation; RM3; ss.
OS Homo sapiens.
FH key Location/Qualifiers
FT CDS 201..1211
FT FT /*tag= a
FT FT /product= "Human RM3 seven transmembrane receptor"
PD 02-JUN-1998.
PE 17-NOV-1993: 153848.
PR 17-NOV-1992: US-977452.
PI (ICOS-) ICOS CORP.
PI Godiska R, Gray PW, Schweickart VL;
PI WPI; 98-332132/29.
DR P-PDB: W48734.
PT DNA encoding V28 seven transmembrane receptor polypeptide - useful
PT for producing recombinant polypeptide and anti-V28 antibodies, and
PT in screening assays for V28 agonists and antagonists
PS Example 11: Columns 89-94; 56bp; English.
CC The present novel sequence represents the human RM3 cDNA isolated
CC from a human macrophage cDNA library. The invention claims for a
CC the full length V28 genomic DNA sequence (V18343) isolated from a human
CC placenta genomic library. The V28 (W48722) and RM3 proteins are seven
CC transmembrane (7TM) receptors which are probably involved in signal
CC transduction. The invention also claims that cells transformed with
CC V28 DNA can be used to produce the recombinant polypeptide, to produce
CC anti-V28 antibodies or in screening assays for V28 agonists or
CC antagonists. The antibodies, agonists and antagonists could then be
CC used to modulate V28 receptor-ligand binding, for e.g. in immunological
CC and/or inflammatory events in vivo.
SQ Sequence 1317 BP: 352 A; 342 C; 265 G; 378 T;

Query Match 61.4%; Score 1067; DB 1; Length 1317;
Best local similarity 100.0%; Pred. No. 0;
Matches 1067; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 ATATACACTTCAGATACTACACCGAGAAATGGCTCAGGGAGACTGATCCATGAAAG 165
DB 171 ATATACACTTCAGATACTACACCGAGAAATGGCTCAGGGAGACTGATCCATGAAAG 230
QY 166 GAACCCGTTCCGTGAAGAAATGCTAATTGATAAATCTTCCGCCACCATCTAC 225
DB 231 GAACCCGTTCCGTGAAGAAATGCTAATTGATAAATCTTCCGCCACCATCTAC 290
QY 226 TCCATCATCTTCTTAAGTGGCATGTGGGCAATGATGTGTCATCCGCGCATGGGTAC 285
DB 291 TCCATCATCTTCTTAAGTGGCATGTGGGCAATGATGTGTCATCCGCGCATGGGTAC 350
QY 286 CAGAGAACTGAGAACATGACGACAGCAAGTACAGGCTGCAGCTGTGAGTGGCACTC 345
DB 351 CAGAGAACTGAGAACATGACGACAGCAAGTACAGGCTGCAGCTGTGAGTGGCACTC 410
QY 346 CTCTTTGTCATCAGGCTCCCTTGGGCGAGTGGAGCCCTGGCAAACTGGTACTTTGGG 405
DB 411 CTCTTTGTCATCAGGCTCCCTTGGGCGAGTGGAGCCCTGGCAAACTGGTACTTTGGG 470
QY 406 AACCTCTATGCAAGGAGTGCATGTCATCTACACAGTCAACCTCTACAGCACTGTCTC 465
DB 471 AACCTCTATGCAAGGAGTGCATGTCATCTACACAGTCAACCTCTACAGCACTGTCTC 530
QY 466 ATCTTGCCCTTATCATGCTGGAACCGCTACCTGCGCATGCTCCAGCCCAACAGTCTAG 525
DB 531 ATCTTGCCCTTATCATGCTGGAACCGCTACCTGCGCATGCTCCAGCCCAACAGTCTAG 590
QY 526 AGGCCAAGGAGTGTGGTGAAGGTGTGTATGTTGGCTGTGATCCTGCGCTC 585

DB 591 AGGCCAAGGAGTGTGGTGAAGGTGTGTATGTTGGCTGTGATCCTGCGCTC 650
QY 586 CTGCTGACTATTTCCGACATTCATCTTTGCCAACGTCAGTGGAGAGATGACATATATC 645
DB 651 CTGCTGACTATTTCCGACATTCATCTTTGCCAACGTCAGTGGAGAGATGACATATATC 710
QY 646 TGTGACCGCTTACCCCAATGACTGTGGGAGTGTGTGTCCAGTTTCAGACATCATG 705
DB 711 TGTGACCGCTTACCCCAATGACTGTGGGAGTGTGTGTCCAGTTTCAGACATCATG 770
QY 706 GTTGCCCTTATTCCTGCTGTGATTTGTCATCTGTCTGTATGATATATATCTTCCAG 765
DB 771 GTTGCCCTTATTCCTGCTGTGATTTGTCATCTGTCTGTATGATATATATCTTCCAG 830
QY 766 CTGTCACTCCCAAGGCGCCACAGAAAGCCCTTAAGACACATCTCTCATC 825
DB 831 CTGTCACTCCCAAGGCGCCACAGAAAGCCCTTAAGACACATCTCTCATC 890
QY 826 CTGCTTCTTCCGCTGTGTGGCTGCTTACTACTCATTTGGATCAGATGATCCTTCATC 885
DB 891 CTGCTTCTTCCGCTGTGTGGCTGCTTACTACTCATTTGGATCAGATGATCCTTCATC 950
QY 886 CTCTGGAATCATCAAGCAAGGCTGTGATTTGAGAACACTGTGCACAAGTGAATTC 945
DB 951 CTCTGGAATCATCAAGCAAGGCTGTGATTTGAGAACACTGTGCACAAGTGAATTC 1010
QY 946 ATACCGAGGCTTACTTCTTCCACTGTGTGTGACATCCCATCTCTATCTTCTT 1005
DB 1011 ATACCGAGGCTTACTTCTTCCACTGTGTGTGACATCCCATCTCTATCTTCTT 1070
QY 1006 GGAGCCAAATTTAAACCTCTGCCAGCAGCATCCTCTGTGAGCAGAGGCTCCAC 1065
DB 1071 GGAGCCAAATTTAAACCTCTGCCAGCAGCATCCTCTGTGAGCAGAGGCTCCAC 1130
QY 1066 CTCAATATCTCTCCAAAGGAGGTGACATCTCTGTTCACGTAGTCTGAG 1125
DB 1131 CTCAATATCTCTCCAAAGGAGGTGACATCTCTGTTCACGTAGTCTGAG 1190

QY 1126 TCTTCAAGTTTCACTCCAGCTACACAGATGATAAGCTTTT 1172
DB 1191 TCTTCAAGTTTCACTCCAGCTACACAGATGATAAGCTTTT 1237

RESULT 6
ID 066179 standard: cDNA; 1317 BP.
AC 066179:
DT 07-FEB-1995 (first entry)
DE Seven transmembrane receptor (RM3) coding sequence.
KW Primer: seven transmembrane receptor; receptor; amplification; PCR;
KW polymerase chain reaction; ss.
OS Homo sapiens.
FH key Location/Qualifiers
FT CDS 201..1214
FT FT /*tag= a
FT FT /product= "seven transmembrane receptor".
PD WO9412635-A.
PE 09-JUN-1994.
PE 17-NOV-1993: U11153.
PR 17-NOV-1992: US-977452.
PI (ICOS-) ICOS CORP.
PI Godiska R, Gray PW, Schweickart VL;
PI WPI; 94-200264/24.
DR P-PDB: R53753.
PT DNA encoding seven transmembrane receptors - used to develop
PT prods. for use as therapeutic or diagnostic agents for conditions
PT involving the receptors.
PS Example 11: Page 82-83; 100bp; English.
CC Two primers (066148, 066149) were used in a PCR reaction containing
CC human macrophage cDNA library in plasmid pRC/CMV. When the PCR
CC products were subjected to agarose gel electrophoresis a faint band
CC of 180-200 base pairs was observed. Re-amplified material was
CC digested with BamHI and HindIII and cloned into the plasmid

CC Bluescript SK-. Of sixteen clones sequenced, two contained a unique
CC sequence termed RM3. Specific primers for the partial RM3 clone were
CC used to identify this full length RM3 cDNA clone.
50 Sequence 1317 BP: 332 A; 343 C; 264 G; 378 T;

Query Match	58.5%	Score 1016;	DB 1;	Length 1317;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1066;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0

QY	105	AATACCTTCAAGTACTACACCGAGAAATGGCTCAGGGAGATATATCTCCATGAG	165
Db	171	ATATACCTTCAAGTACTACACCGAGAAATGGCTCAGGGAGATATATCTCCATGAG	230
QY	166	GAACCCGTGTTCCGTGAGAAATGCTAAATTTCAATAAATCTTGCTGCCACATATC	225
Db	231	GAACCCGTGTTCCGTGAGAAATGCTAAATTTCAATAAATCTTGCTGCCACATATC	290
QY	226	TCCATCATCTTTTAACTGGCAATTTGGGCAATGCAATGGTATCTCTGGTCAATG	285
Db	291	TCCATCATCTTTTAACTGGCAATTTGGGCAATGCAATGGTATCTCTGGTCAATG	350
QY	286	CAGAGAACTGAAAGCATGACGAGCAAGTACAGCTCAGCTCAGTGGCGACCTC	345
Db	351	CAGAGAACTGAAAGCATGACGAGCAAGTACAGCTCAGCTCAGTGGCGACCTC	410
QY	346	CTCTTTGTATCAAGCTTCCCTCTTGAGGAGATGATGAGCCGTGGCAACAGTGTATG	405
Db	411	CTCTTTGTATCAAGCTTCCCTCTTGAGGAGATGATGAGCCGTGGCAACAGTGTATG	470
QY	406	AACCTTCATGCAAGGCAGTCCATGTATCTACACAGTAACTCTACAGCAGTCTC	465
Db	471	AACCTTCATGCAAGGCAGTCCATGTATCTACACAGTAACTCTACAGCAGTCTC	530
QY	465	ATCCGCGCTTCAATGATGTGGAGCCGTCACCGGCAATGTCACAGCCACCAACAGTAG	525
Db	531	ATCCGCGCTTCAATGATGTGGAGCCGTCACCGGCAATGTCACAGCCACCAACAGTAG	590
QY	526	AGGCCAAGAAAGCTGTGGCTGAAAGAGTGATATGTTGGCGTGTGGATCCCTCCCTC	585
Db	591	AGGCCAAGAAAGCTGTGGCTGAAAGAGTGATATGTTGGCGTGTGGATCCCTCCCTC	650
QY	586	CTGCTGACTATTCGCACTTCACTTTTGCCAAAGTCAGTGAAGCAGATGACAGATATATC	645
Db	651	CTGCTGACTATTCGCACTTCACTTTTGCCAAAGTCAGTGAAGCAGATGACAGATATATC	710
QY	645	TGTGACCGCTTACCCCAATGACTGTGGGTGTGTGTGTCATTTGACAGCATATG	705
Db	711	TGTGACCGCTTACCCCAATGACTGTGGGTGTGTGTGTCATTTGACAGCATATG	770
QY	706	GTTGGCCTTATCTGCTGGTATGTATCTATCTGCTGCTGATGTGATATCATCTCCAG	765
Db	771	GTTGGCCTTATCTGCTGGTATGTATCTATCTGCTGCTGATGTGATATCATCTCCAG	830
QY	766	CTGTACACTTCACAGGGCCACCAAGAGCCCAAGGCCCTTAAAGACACAGTATCTCTATC	825
Db	831	CTGTACACTTCACAGGGCCACCAAGAGCCCAAGGCCCTTAAAGACACAGTATCTCTATC	890
QY	826	CTGGGCTTCTGGCCGTGTGGCTGTGCTTACTATCATTTGGATGAGATGAGACTCTTATC	885
Db	891	CTGGGCTTCTGGCCGTGTGGCTGTGCTTACTATCATTTGGATGAGATGAGACTCTTATC	950
QY	886	CTCTCTGAAATCATCAAGCAAGGGTGTGAGTTTGAAGAACTGTGCACAGTGGATTTC	945
Db	951	CTCTCTGAAATCATCAAGCAAGGGTGTGAGTTTGAAGAACTGTGTACAAAGTGGATTTC	1010
QY	946	ATACCGAGGCCCTAGCTTTCTCACTGTGTCTGAACCCCATCTTATGTTTCCTT	1005
Db	1011	ATACCGAGGCCCTAGCTTTCTCACTGTGTCTGAACCCCATCTTATGTTTCCTT	1070
QY	1006	GGAGCCAAATTTAAACCTCTGCCAGCACAGCACTACTCTGTGTAGCAGAGGTCCAGC	1065
Db	1071	GGAGCCAAATTTAAACCTCTGCCAGCACAGCACTACTCTGTGTAGCAGAGGTCCAGC	1130

Qy	1086	CTCAAGATCCCTCTCCAAAGGAAGGCGAGGTGACATCATCTGTTTCCACGTAGCTGAG	1125
Db	1131	CTCAAGATCCCTCTCCAAAGGAAGGCGAGGTGACATCATCTGTTTCCACGTAGCTGAG	1150
Qy	1126	TCTTCAGTTTCTACCTCCAGCTACACAGATGTAAAGACCTTTTT	1172
Db	1191	TCTTCAGTTTCTACCTCCAGCTACACAGATGTAAAGACCTTTTT	1237

[illegible]

RESULT 7

ID	Accession	Standard	CDNA	611 BP
Y15883				
AC	X15883			
DT	12-MAY-1999	(first entry)		
DE	CDNA encoding a partial CXCR4B protein.			
KW	G-protein coupled receptor; CXCR4B; human; splice variant;			
KW	chemokine receptor; CXCR4; vaccine; gene therapy; HIV-1; HIV-1;			
KW	cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;			
KW	Parkinson's disease; acute heart failure; hypotension; hypertension;			
KW	urinary retention; osteoporosis; angina pectoris; myocardial infarction;			
KW	stroke; ulcer; allergy; benign prostatic hypertrophy; migraine;			
KW	vomiting; psychotic disorder; neurological disorder; anxiety;			
KW	schizophrenia; manic depression; delirium; dementia; mental retardation;			
KW	dyskinesia; Huntingtons disease; Gilles de la Tourette syndrome; ss.			
OS	Homo sapiens.			
PN	EP-897860-A2.			
PD	24-FEB-1999.			
PF	07-AUG-1998; 306324.			
PR	24-JUL-1998; US-056601.			
PR	20-AUG-1997; US-056601.			
PA	(SMK) SMITHKLINE BEECHAM CORP.			
PI	Gupta SK, Pillarsetti K;			
DR	WPI: 99-134643/12.			
DR	P-P5DB: W97363.			
PT	New G protein coupled receptor (CXCR4B) polypeptide and			
PT	polynucleotide, human splice variant of a chemokine receptor -			
PT	useful as diagnostic reagents and for prevention and treatment of			
PT	HIV infection, cancer, stroke and dementia			
PS	Claim 13; Page 18; 24pp; English.			
CC	The present sequence encodes a partial G-protein coupled receptor			
CC	designated CXCR4B, which is a human splice variant of a chemokine			
CC	receptor CXCR4. CXCR4B polynucleotides and polypeptides are useful for			
CC	diagnosing susceptibility to diseases by detecting mutations or			
CC	polymorphisms in the CXCR4B gene or analysing for the presence or amount			
CC	of CXCR4B polypeptide expressed in a patient. CXCR4B polypeptides and			
CC	polynucleotides are also useful for screening for antagonists and			
CC	agonists which can be used to treat conditions associated with CXCR4B			
CC	polypeptide imbalance. CXCR4B polypeptides can be administered directly			
CC	(as a vaccine) or via a vector (gene therapy) to prevent disease.			
CC	Diseases diagnosed, prevented and treated include: HIV-1 and HIV-1			
CC	infections; cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;			
CC	Parkinson's disease; acute heart failure; hypotension; hypertension;			
CC	urinary retention; osteoporosis; angina pectoris; myocardial infarction;			
CC	stroke; ulcers; allergies; benign prostatic hypertrophy; migraine;			
CC	vomiting; psychotic and neurological disorders, including anxiety,			
CC	schizophrenia, manic depression, delirium, dementia and severe mental			
CC	retardation; and dyskinesias, such as Huntingtons disease or Gilles de			
CC	la Tourette syndrome. CXCR4B polypeptides are useful for mapping genes			
CC	to chromosomes, allowing gene inheritance to be studied through linkage			
CC	analysis.			
SO	Sequence	611 BP;	166 A;	125 C; 139 G; 181 T;
QY	Query Match	14.3%;	Score 249;	DB 1; Length 611;
DB	Best Local Similarity	100.0%;	Pred. No. 3e-80;	
Y	Matches 249;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Y	166 GAACCTCTTTCGATGAGAAATATCTTAATTCAATTAATAATTTCTGCGCCATCTAC			225


```

Db 423 GAACCTGTTTCGGAGAAAATGTAATTTCAATTAATTTCTCTGCCACCACTTAC 482
Oy 226 TCCATCATCTTCTTAAGTGGCAATGTTGGCAATGATTCCTGTCATGGTTAC 285
Db 483 TCCATCATCTTCTTAAGTGGCAATGTTGGCAATGATTCCTGTCATGGTTAC 542
Oy 286 CAGAGAACTAGAGAGCATGACGACAGTACAGCTCAGCTGACCTGACGCGACCTC 345
Db 543 CAGAGAACTAGAGAGCATGACGACAGTACAGCTCAGCTGACGCGACCTC 602
Oy 346 CTCTTTGTC 354
Db 603 CTCTTTGTC 611

RESULT 8
T20146
ID T20146 standard; cDNA to mRNA; 218 BP.
AC T20146:
DE 31-JUL-1996 (first entry)
DE Human gene signature H0MGS01292.
DE Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KM human; cloning; mapping; non-biased library; diagnosis; detection;
KM cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATSU) MATSUBARA K.
PA (OKUBO) OKUBO K.
PI Matsubara K, Okubo K;
PI WPI: 95-206931/27.
PI Identifying gene signatures in 3'-directed human cDNA library - e.g.
PI for diagnosis of abnormal cell function, by preparing cDNA that
PI reflects relative abundance of corresp. mRNA in specific human
PI tissues
PS Claim 1: Page 571-572; 2245pp: Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in 119001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues: synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 218 BP; 61 A; 38 C; 44 G; 75 T;

Query Match 12.6%; Score 218; DB 1; Length 218;
Best Local Similarity 100.0%; Pred. No. 3.9e-69; Indels 0; Gaps 0;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1634 GATTAAGTTGTAATAAAGTACATGTTAACTTAA 1671
Db 181 GATTAAGTTGTAATAAAGTACATGTTAACTTAA 218

RESULT 9
O44391
ID O44391 standard; cDNA to mRNA; 3581 BP.
AC O44391:
DE 14-SEP-1994 (first entry)
DE Sequence of murine OSF-4 cDNA.
DE OSF-4; cadherin; growth factor; osteogenesis; osteoblast; therapy;
KM diagnosis; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT cds 284..2671
FT /*tag= a

PN EP-585801-A.
PD 09-MAR-1994.
PF 25-AUG-1993; 113602.
PR 28-AUG-1992; JP-230028.
PA (FARH) HOECHST JAPAN LTD.
PI Amann F, Kawai S, Okazaki M, Takeshita S, Tsujimura A;
PI WPI: 94-076152/10.
PI P-PSDB: R49730.
PI New bone related, cadherin-like OSF-4 proteins - for treatment
PI and diagnosis of bone metabolic disease, and nucleic acid
PI encoding them
PS Claim 3: Page 13-17; 34pp: English.
CC cDNA libraries were constructed from the mouse osteoblastic cell
CC line MC3T3-E1 and from mouse liver tissue, amplified by PCR, and
CC then as much common DNA as possible removed by hybridisation between
CC the 2 libraries. Residual E1-specific DNA was amplified, inserted
CC into lambda gt10 and screened by plaque hybridisation. A minibank of
CC 273 E-specific clones was recovered, their inserts amplified and
CC used to screen total RNA from both cell types. One clone specific
CC for E1 was identified and sequenced. The insert from this clone was
CC used to screen cDNA prepd. from E1 RNA and the longest posn. insert
CC cloned in pGEM 112f (+) to give pKOT164. This insert was sequenced;
CC it encoded the 796 AA mouse precursor protein (O44391/R49730). The
CC insert was also used to screen a cDNA bank prepd. from human
CC osteosarcoma to identify 2 clones encoding the 2 human precursor
CC proteins - OSF-4-1 and OSF-4-2 (O44392/R49731 and O44393/R49732
CC respectively).
SQ Sequence 3581 BP; 1121 A; 774 C; 800 G; 886 T;

Query Match 3.1%; Score 53; DB 1; Length 3581;
Best Local Similarity 100.0%; Pred. No. 9.3e-11; Indels 0; Gaps 0;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1669 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCG 1721
Db 3523 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCG 3575

RESULT 10
O65607
ID O65607 standard; cDNA; 1338 BP.
AC O65607:
DE 31-JAN-1995 (first entry)
DE Rabbit zona pellucida ZPC coding sequence.
DE Rabbit; lapine; zona pellucida; ZPC; immunoneutralization; ds.
OS Oryctolagus cuniculus.
FH Key Location/Qualifiers
FT cds 17..1264
FT /*tag= a
FT /product= ZPC

PN WO9411019-A.
PD 26-MAY-1994.
PF 06-NOV-1993; U10851.
PR 09-NOV-1992; US-973341.
PR 29-JAN-1993; US-012990.

```



```
PA (ZONA-) ZONAGEN INC.
PI Harris JD, Hsu KT, Podolski JS;
DR WPI: 94-183156/22.
DR P-PSDB: R55197.
PI Use of zona pellucida proteins and antibodies - for inducing
PT reproducible transient infertility or permanent sterility in
PT female mammals
PI Claim 22: Page 81-83; 154pp; English.
CC A cDNA library was prepared from mRNA isolated from ovaries removed
CC from 5 week old rabbits. The lambda gt10 library was screened with
CC a porcine ZPC cDNA probe (Q055606). Two positive clones were sequenced
CC and were found to be identical except that one contained 4 extra
CC nucleotides at the 5'-end. The determined sequence was 75%
CC homologous to DNA coding for porcine ZPC.
SO Sequence 1338 BP; 255 A; 445 C; 402 G; 236 T;

Query Match 3.1%; Score 53; DB 1; Length 1338;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1659 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 1721
DB 1280 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 1332

RESULT 11
V05740
ID V05740 standard; cDNA: 87 BP.
AC V05740.
DT 05-JUN-1998 (first entry)
DE Nucleotide sequence of the 3' portion of clone AS301_2.
KW Secreted protein; homology; antibody; immunoassay reagent;
KW nutritional supplement; therapeutic activity; ds.
OS Homo sapiens.
PN WO9746683-A2.
PD 11-DEC-1997.
PF 06-JUN-1997; 009878.
PR 07-JUN-1996; US-659224.
PA (GENY ) GENETICS INST INC.
PI Bowman M, Evans C, Jacobs K, Lavallic ER, McCly JM,
PI Merberg D, Racie LA, Spaulding V, Treacy M;
DR WPI: 98-042191/04.
PT Nucleic acids encoding secreted proteins from clones within ATCC
PT 98076 - useful as immuno-modulators, anti-proliferative agents,
PT regulators of cell differentiation and tissue growth, etc
PS Claim 40: Page 76; 99pp; English.
CC The present sequence represents the nucleotide sequence of the 3'
CC portion of clone AS301_2. The clone was isolated from a human fetal
CC brain cDNA library using probe V05760. AS301_2 is a full length clone
CC encoding a secreted protein. V05739 provides 5' portion of the nucleotide
CC sequence, and V05738 an internal sequence. The nucleic acid can be used
CC for expression of recombinant proteins, as tissue, molecular weight or
CC chromosome markers, indicators of genetic disorders and sources
CC of probes and primers. They can also be used to generate anti-protein
CC or anti-DNA antibodies and as components of interaction trap assays etc.
CC The protein is useful for raising antibodies, as immunoassay reagents
CC and as nutritional supplements. The protein may possibly have any of a
CC great variety of therapeutic activities.
SO Sequence 87 BP; 79 A; 4 C; 4 G; 0 U;
```

```
ID V05728 standard; cDNA: 94 BP.
AC V05728.
DT 05-JUN-1998 (first entry)
DE Nucleotide sequence of the 3' portion from clone A1147_1.
KW Secreted protein; antibody; immunoassay reagent;
KW nutritional supplement; therapeutic activity; murine; calmegin; ds.
OS Homo sapiens.
PN WO9746683-A2.
PD 11-DEC-1997.
PF 06-JUN-1997; 009878.
PR 07-JUN-1996; US-659224.
PA (GENY ) GENETICS INST INC.
PI Bowman M, Evans C, Jacobs K, Lavallic ER, McCly JM,
PI Merberg D, Racie LA, Spaulding V, Treacy M;
DR WPI: 98-042191/04.
PT Nucleic acids encoding secreted proteins from clones within ATCC
PT 98076 - useful as immuno-modulators, anti-proliferative agents,
PT regulators of cell differentiation and tissue growth, etc
PS Claim 25: Page 67; 99pp; English.
CC The present sequence represents the nucleotide sequence of the 3'
CC portion of clone A1147_1. The clone was isolated from a human adult
CC testes cDNA library using probe V05755. A1147_1 is a full length
CC clone encoding a secreted protein. V05727 provides the 5' portion
CC sequence. A1147_1 shows some homology with murine calmegin, a
CC Ca2+-binding protein that is specifically expressed in spermatogenesis.
CC A1147_1 may share some of its activity. The nucleic acid can be used for
CC expression of recombinant proteins, as tissue, molecular weight or
CC chromosome markers, indicators of genetic disorders and sources of probes
CC and primers. They can also be used to generate anti-protein or anti-DNA
CC antibodies and as components of interaction trap assays etc. The protein
CC is useful for raising antibodies, as immunoassay reagents and as
CC nutritional supplements. The protein may possibly have any of a great
CC variety of therapeutic activities.
SO Sequence 94 BP; 86 A; 4 C; 4 G; 0 U;

Query Match 3.1%; Score 53; DB 1; Length 94;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1659 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 1721
DB 42 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 94

RESULT 13
V61487
ID V61487 standard; cDNA: 4237 BP.
AC V61487.
DT 11-JAN-1999 (first entry)
DE Human secreted protein fe366.1 cDNA.
KW Secreted protein; human; fe366_1; ds.
OS Homo sapiens.
PN key location/Qualifiers
FT CDS 3746..4030
FT /tag= a
FT WO9841539-A2.
PD 24-SEP-1998.
PF 19-MAR-1998; 005474.
PR 18-MAR-1998; US-040963.
PA 19-MAR-1997; US-820493.
PA (GENY ) GENETICS INST INC.
PI Agostino M, Jacobs K, Lavallic ER, McCoy JM, Merberg D,
PI Racie LA, Spaulding V, Treacy M;
DR WPI: 98-521163/44.
DR P-PSDB: W79096.
PI New polynucleotide(s) encoding secreted human proteins - derived
PI from human foetal kidney, adult testes and adult or foetal brain
PI cDNA libraries
PS Claim 36: Page 88-91; 112pp; English.
CC This full-length cDNA clone, designated fe366.1, codes for a novel
CC secreted human protein (see W79096). It was isolated from a human
CC adult brain cDNA library using methods which are selective for
```

```
CC cDNAs encoding secreted proteins, or was identified as encoding
CC a secreted or transmembrane protein on the basis of computer
CC analysis of the encoding protein. The nucleotide sequence shows
CC homology to some database sequences, and may contain a CAA repeat
CC and/or Alu repetitive element. The invention provides cDNA clones
CC (see V61477-87) from human foetal kidney, adult testis, and adult
CC or foetal brain cDNA libraries that code for secreted proteins
CC (see W79087-97). These clones are deposited as ATCC 98364. The
CC polynucleotides and proteins are predicted to have useful
CC biological activities which would make them suitable for treating,
CC preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested
CC activities include nutritional, immune stimulating (e.g. as
CC vaccines) or suppressing, haematopoiesis regulating, tissue growth,
CC activin/inhibin, chemotactic/chemokinetic, haemostatic and
CC thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour
CC invasion suppressor and tumour inhibition activities. The
CC polynucleotides are also stated to be useful for gene therapy.
SQ Sequence 4237 BP; 1330 A; 778 C; 784 G; 1338 T;

Query Match
Best Local Similarity 100.0%; Score 53; DB 1; Length 4237;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1669 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCGC 1721
DB 2554 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCGC 2606

RESULT 14
V64789
ID V64789 standard; CDNA; 1338 BP.
AC V64789;
DT 29-JAN-1999 (first entry)
DE Rabbit ZPC CDNA.
KW ZPC; zona pellucida; infertility; sterility; immunocontraceptive;
OS Oryctolagus cuniculus;
FH Key Location/Qualifiers
FT CDS 17..1264
FT /tag= a
FT /product= "ZPC"

PN US5837497-A.
PD 17-NOV-1998.
PF 07-JUN-1995; 484993.
PR 09-NOV-1993; US-149223.
PR 09-NOV-1992; US-973441.
PR 29-JAN-1993; US-012990.
PR 07-JUN-1995; US-484993.
PA (ZONA-) ZONAGEN INC.
PI Harris JD;
DR WPI: 99-023447/02.
PT P-SPB; W81807.
PT Isolated zona pellucida DNA from different mammals - used to develop
PT products which can be used for vaccination to induce transient
PT infertility or permanent sterility in female mammals
PS Claim 1: Column 59-62; 84pp; English.
CC This sequence encodes a rabbit ZPC protein isolated from zona pellucida.
CC This protein can be used in a method for specifically inducing transient
CC infertility or permanent sterility in a host animal by selective
CC vaccination with specific zona pellucida proteins or
CC immunocontraceptively active fragments.
SQ Sequence 1338 BP; 255 A; 445 C; 402 G; 236 T;
```

```
RESULT 15
V84799
ID V84799 standard; CDNA; 2744 BP.
AC V84799;
DT 30-MAR-1999 (first entry)
DE Nucleotide sequence encoding human prohormone convertase 4.
KW ds: human; prohormone convertase; testicular prohormone; testicular cell;
KW fertility; spermatogenesis; embryo development; chromosome 19; 19p13.3.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 61..2328
FT /tag= a
FT /product= "prohormone convertase 4"

PN M09650560-A1.
PD 12-NOV-1998.
PF 01-MAY-1998; U08871.
PR 06-MAY-1997; US-044015.
PA (ZYMO ) ZYMOGENETICS INC.
PI Jaspers SR, Lok S;
DR WPI: 99-059673/05.
DR P-PSDB; W81365.
PT New nucleic acid encoding human prohormone convertase 4 - useful
PT for, e.g. identifying modulators and new testicular prohormones, and
PT diagnosing chromosomal abnormalities
PS Claim 2: Page 70-75; 95pp; English.
CC Human prohormone convertase 4 can be used to raise antibodies; for in
CC vitro identification of modulators and to identify, or determine
CC function of, (new) testicular prohormones, and for processing of such
CC prohormones (e.g. for stimulating proliferation or differentiation of
CC testicular cells). mRNA for PC4 is detected only in testis, suggesting
CC its involvement in fertility, spermatogenesis and early embryo
CC development. The nucleic acid and its fragments (particularly probes and
CC primers) are used to detect abnormalities in chromosome 19 (the gene for
CC PC4 is located at 19p13.3).
SQ Sequence 2744 BP; 552 A; 961 C; 833 G; 398 T;
```

```
Query Match
Best Local Similarity 100.0%; Score 53; DB 1; Length 2744;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1669 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCGC 1721
DB 2692 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCCGC 2744

RESULT 16
Q78635/C
ID Q78635 standard; DNA; 61 BP.
AC Q78635;
DT 14-JUL-1995 (first entry)
DE Human chondromodulin-1 cDNA primer.
KW Chondromodulin-1; chondrocyte growth regulation;
KW vascular endothelial cell growth; primer; ss.
OS Synthetic.
PN EP-624645-A.
PD 17-NOV-1994.
PF 11-MAY-1994; 107364.
PR 11-MAY-1993; JP-109620.
PR 17-DEC-1993; JP-318298.
PA (SUZUKI) SUZUKI F.
PA (MITU) MITSUBISHI KASEI CORP.
PI Hiraki Y, Kohara A, Kondo J, Mori A, Suzuki F;
PI Suzuki Y, Takahashi K, Yamada E;
DR WPI: 94-350785/44.
PT Human chondromodulin-1 protein - stimulates growth of
PT chondrocytes and inhibits vascular endothelial cell growth
PS Example 4: Page 10; 34pp; English.
CC Q78635 and Q78636 are a pair of primers for Q78627-Q78632 which
CC encode R65486-R65491 respectively, human chondromodulin-1 (CM-1),
CC a chondrocyte growth regulator. CM-1 as part of a pharmaceutical
```

CC compen. with a carrier, excipient or solvent, can be used to
CC stimulate chondrocyte growth. CM-1 also inhibits vascular
CC endothelial cell growth, by promoting the differential
CC potency of chondrocytes.
SQ Sequence 61 BP; 2 A; 8 C; 8 G; 43 T;

Query Match 2.9%; Score 51; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 1,1e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1673 AAAAAAAAAAAAAAAAAAAAAAAAAAGCGCCGCCA 1723
DB 61 AAAAAAAAAAAAAAAAAAAAAAAAAAGCGCCGCCA 11

RESULT 17

X26148 17
ID X26148 standard; CDNA: 775 BP.
AC X26148;
DT 21-MAY-1999 (first entry)
DE Human pan-s/tk-1A receptor kinase encoding cDNA.
KW pancreatic serine/threonine kinase receptor; pan-s/tk; cell growth;
KW differentiation; gene therapy; ligand; screening; cell proliferation;
KW regulator; pancreatic; ds.
OS Homo sapiens.
PN MO9907854-A2.
PD 18-FEB-1999.
PE 11-AUG-1997; US-909354.
PR 11-AUG-1997; US-909354.
PT (ONTO-) ONTOGENY INC.
PI Barker DD, Miao N, Pang K;
DR WPI: 99-167430/14.
P-PSDB: Y01496.
PT New pancreatic serine/threonine kinase (pan-s/tk) receptor
PT polypeptides - useful for modulating cell proliferation,
PT differentiation and survival in animals
PS Claim 28: Page 66: 77pp; English.
CC The invention relates to isolated and/or recombinant pancreatic serine/
CC threonine kinase receptor (pan-s/tk) polypeptides. Sequences of rat and
CC human pan-s/tk polypeptides and nucleic acids encoding them are provided.
CC Host cells containing vectors comprising the pan-s/tk genes are
CC used for the recombinant expression of the proteins. The pan-s/tk
CC polypeptides are useful for modulating cell growth, differentiation or
CC survival in an animal cell. The polynucleotides are also useful for
CC detecting pan-s/tk ligand on cells in samples and for gene therapy.
CC Ligands identified, together with soluble polypeptides, are useful for
CC screening test compounds that modulate the bioactivity of a pan-s/tk
CC receptor. The discovery of this new class of pan-s/tk polypeptides
CC provides an alternative method of mediating cell proliferation,
CC differentiation and survival in animals. The polypeptides also represent
CC the first identified specific regulators of pancreatic development. The
CC present sequence represents a cDNA encoding a human pan-s/tk-1A receptor
CC polypeptide.
SQ Sequence 775 BP; 231 A; 188 C; 179 G; 177 T;

Query Match 2.9%; Score 51; DB 1; Length 775;
Best Local Similarity 100.0%; Pred. No. 6,6e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1663 AACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
DB 716 AACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 766

RESULT 18
X26151 18
ID X26151 standard; CDNA: 1718 BP.
AC X26151;
DT 21-MAY-1999 (first entry)
DE 3' cDNA sequence of human pan-s/tk-2 receptor polypeptide.
KW Pancreatic serine/threonine kinase receptor; pan-s/tk; cell growth;

KW differentiation; gene therapy; ligand; screening; cell proliferation;
KW regulator; pancreatic; ds.
OS Homo sapiens.
PN MO9907854-A2.
PD 18-FEB-1999.
PE 11-AUG-1998; U16640.
PR 11-AUG-1997; US-909354.
PT (ONTO-) ONTOGENY INC.
PI Barker DD, Miao N, Pang K;
DR WPI: 99-167430/14.

PT New pancreatic serine/threonine kinase (pan-s/tk) receptor
PT polypeptides - useful for modulating cell proliferation,
PT differentiation and survival in animals
PS Claim 28: Page 73: 77pp; English.
CC The invention relates to isolated and/or recombinant pancreatic serine/
CC threonine kinase receptor (pan-s/tk) polypeptides. Sequences of rat and
CC human pan-s/tk polypeptides and nucleic acids encoding them are provided.
CC Host cells containing vectors comprising the pan-s/tk genes are
CC used for the recombinant expression of the proteins. The pan-s/tk
CC polypeptides are useful for modulating cell growth, differentiation or
CC survival in an animal cell. The polynucleotides are also useful for
CC detecting pan-s/tk ligand on cells in samples and for gene therapy.
CC Ligands identified, together with soluble polypeptides, are useful for
CC screening test compounds that modulate the bioactivity of a pan-s/tk
CC receptor. The discovery of this new class of pan-s/tk polypeptides
CC provides an alternative method of mediating cell proliferation,
CC differentiation and survival in animals. The polypeptides also represent
CC the first identified specific regulators of pancreatic development. The
CC present sequence represents the 3' cDNA sequence of a human pan-s/tk-2
CC receptor polypeptide.
SQ Sequence 1718 BP; 495 A; 335 C; 394 G; 492 T;

Query Match 2.9%; Score 51; DB 1; Length 1718;
Best Local Similarity 100.0%; Pred. No. 5,5e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1663 AACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
DB 1659 AACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1709

RESULT 19

T74484 19
ID T74484 standard; CDNA: 495 BP.
AC T74484;
DT 01-NOV-1997 (first entry)
DE Pilosulin 2 precursor cDNA clone Myr p2.
KW Pilosulin 2; Australian jumper ant; venom; allergen; hypotensive;
KW blood pressure; cell growth; ss.
OS Myrmecia pilosula.
FH Key Location/Qualifiers
FT met_peptide 44..271
FT /tag= a
FT signal_peptide 44..121
FT /*tag= b
FT mat_peptide 188..268
FT /*tag= c
FT /product= pilosulin A
PN MO9713854-A1.
PD 17-APR-1997.
PE 09-OCT-1996; AU0632.
PR 09-OCT-1995; AU-005839.
PA (NSYD-) NORTHERN SYDNEY AREA HEALTH SERVICE.
PI Baldo B, Donovan G;
DR WPI: 97-235890/21.
DR P-PSDB: W21784.
PT New isolated Pilosulin 2 polypeptide - obtained from Myrmecia
PT Pilosula ant venom, used for reducing blood pressure in animals or
PT for inhibiting cell growth
PS Claim 1: Fig 1: 27pp; English.
CC This sequence is an isolated cDNA clone, designated Myr p2, which
CC includes a sequence encoding a polypeptide that has the biological

CC activity of pilosulin 2 having the amino acid sequence from
 CC residues 49-75 of a 75-amino acid sequence (W21784). It was
 CC isolated from a cDNA library prepared from the abdomens of Myrmecia
 CC pilosula ants by immunoscreening with sera from venom-allergic
 CC individuals. The isolated sequence can be incorporated into an
 CC expression vector and used to produce pilosulin 2 in transformed
 CC host cells. Pilosulin 2 and its analogues or fragments can be used
 CC in claimed methods for reducing the blood pressure of an animal,
 CC and for inhibiting the growth of (killing) cells.
 SO Sequence 495 BP; 201 A; 68 C; 99 G; 127 T;

Query Match 2.9%; Score 50; DB 1; Length 495;
 Best Local Similarity 100.0%; Pred. No. 1.6e-09;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1672 AAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 1721
 DB 446 AAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 495

RESULT 20
 ID V00418 standard; cDNA: 201 BP.

AC V00418; 1998 (first entry)
 DE 3' fragment of clone M8_2.
 KW Human; secreted protein; molecular weight marker; genetic fingerprinting;
 KM antibody production; nutritional supplement; therapy; neutral tissue;
 KW glioblastoma line T9985; clone M8_2; ds.
 OS Homo sapiens.
 PN M09740069A2.
 PD 30-OCT-1997.
 PF 14-APR-1997; 006134.
 PR 19-APR-1996; US-635311.
 PA (GENY) GENETICS INST INC.
 PI Jacobs K, Lavalie ER, McCoy JM, Metberg D, Racie LA,
 PI Spaulding V.
 DR WPI: 97-353776/49.
 PT Isolated nucleic acid clones from ATCC 98028 encode novel secreted
 PT proteins - having many potential uses, e.g. as immunomodulators,
 PT cell proliferation or differentiation inhibitors or hematopoiesis
 PT regulators
 PS Claim 22: Page 69: 114pp; English.
 CC This sequence represents the 3' end of clone M8_2, which is a
 CC polynucleotide of the invention. This sequence was isolated from a human
 CC neural tissue (glioblastoma line T9985) cDNA library. The polynucleotide,
 CC which encodes a secreted protein, can be used, e.g. as a tissue or
 CC molecular weight marker, in genetic fingerprinting, to raise anti-protein
 CC or anti-DNA antibodies and in interaction trap assays. The protein can be
 CC used to assay biological activity, raise antibodies for use in
 CC immunosays, as a marker, to identify inhibitors of its interactions and
 CC as a nutritional supplement. It may also have a very wide range of
 CC therapeutic and biological activities (no examples are given to support
 CC this), e.g. cytokine or modulator of cell proliferation and
 CC differentiation, immunostimulant or immunosuppressant, hematopoiesis
 CC regulator, bone, cartilage, tendon, ligament and/or nerve tissue growth
 CC stimulator, follicle inhibitor/stimulator, chemotactic/chemokinetic,
 CC haemostatic, thrombolytic or anti-inflammatory agent, antimicrobial,
 CC biohythm, metabolism or behaviour modifier, anti-depressant or analgesic
 CC or psoriasis treatative.
 SO Sequence 201 BP; 90 A; 16 C; 43 G; 39 T;

Query Match 2.9%; Score 50; DB 1; Length 201;
 Best Local Similarity 100.0%; Pred. No. 2e-09;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1672 AAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 1721
 DB 152 AAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCGC 201

RESULT 21

ID V46370 standard; cDNA to mRNA; 1877 BP.
 AC V46370;
 DT 20-NOV-1998 (first entry)
 DE Nucleic acid encoding a murine CXC chemokine receptor.
 KW Mouse; CXC chemokine receptor; pre-B cell line DM34;
 KW CXC chemokine pre-B cell stimulatory factor PBSF/SDF-1;
 KW HIV infection; screening; inhibitor; AIDS; ds.
 OS Mus sp.
 PN M09835035-A1.
 PD 13-AUG-1998.
 PF 07-FEB-1997; J00299.
 PR 07-FEB-1997; WO-J00299.
 PA (SHIO) SHIONOGI & CO LTD.
 PI Iizasa H, Kishimoto T, Nagasawa T, Nakajima T, Tachibana K,
 PI Yoshida N, Yoshie O;
 DR WPI: 98-447232/38.
 DR P-PSDB: M64778.

DR Mouse CXC chemokine receptor binding to PBSF/SDF-1 pre-B cell
 PT stimulatory factor. Is useful for screening of potential HIV
 PT infection and AIDS inhibitors
 PT Claim 3: Pages 39-42; 76pp; Japanese.
 CC The present sequence encodes a murine CXC chemokine receptor which
 CC binds to the mouse CXC chemokine pre-B cell stimulatory factor
 CC PBSF/SDF-1. The nucleic acid is isolated from mouse pre-B cell
 CC line DM34. The receptor and cells expressing it can be used in the
 CC study and mapping of the mechanism of HIV infection and in screening
 CC of potential inhibitors of HIV infection and the development of AIDS.
 SO Sequence 1877 BP; 469 A; 451 C; 406 G; 551 T;

Query Match 2.9%; Score 50; DB 1; Length 1877;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1662 TAAACTTAAAAAAAAAAAAAAAAAAAAAAAAA 1711
 DB 1828 TAAACTTAAAAAAAAAAAAAAAAAAAAAAAAA 1877

RESULT 22
 ID 052732 standard; DNA; 56 BP.
 AC 052732;
 DT 20-JUN-1994 (first entry)
 DE Sequence of oligo nucleotide adaptor used for the synthesis of
 DE asymmetrically tailed plasmid primers for use in cloning libraries.
 KW Asymmetrically tailed plasmid primer; adaptor; ds.
 OS Synthetic.
 FH Key
 FT misc-feature Location/Qualifiers
 FT complement (1..4) /tag- a
 FT /label= Sticky end
 FT misc-feature /tag- b
 FT /label= Sticky end
 FT /label= Sticky end
 PN US5270185-A.
 PD 14-DEC-1993. 341523.
 PF 21-APR-1989; 341523.
 PR 21-APR-1989; US-341523.
 PR 12-AUG-1992; US-928856.
 PA (HOFF) HOFFMANN LA ROCHE INC.
 PI Margolske RF;
 DR WPI: 93-404926/50.
 DR Asymmetrically tailed plasmid primers - comprises cut,
 PT double-stranded DNA plasmid, useful for producing cloning
 PT libraries
 PS Claim 4; Figure 3A: 13pp; English.
 CC The inventors claim an asymmetrically tailed plasmid primer

CC comprising cut-ds DNA plasmid contg a functional origin of
CC replication and at least one functional selection marker gene, one 3'
CC terminus of which contains an oligo (dd) extension, the other 3'
CC terminus of which contains an oligo (dc) or oligo (dd) extension
CC terminated by a 3' phosphate group. More specifically, the
CC asymmetrically tailed plasmid primer is produced by SacI and KpnI
CC cleavage of plasmid pIP, and each terminus produced by the cleavage
CC is joined to the complementary adaptor, which is Q52731 or Q52732.
CC Sequence 556 BP: 1 A: 6 G: 8 G: 41 T:

Query Match	2.8%	Score 49;	DB 1;	Length 56;
Best Local Similarity	100.0%	Pred. No. 5.8e-09;		
Matches	49;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0

Qy	1674	AAAAAAAAAAAAAAAAAAAAAAAAAGCGGCGCC	1722
Db	56	AAAAAAAAAAAAAAAAAAAAAAAAAGCGGCGCC	8

Db 56 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCCGCC 8

```

RESULT 23
052734/c
ID 052734 standard; DNA; 56 BP.
AC 052734;
DT 20-JUN-1994 (first entry)
DE Sequence of oligo nucleotide adaptor used for the synthesis of
DE asymmetrically tailed plasmid primers for use in cloning libraries
KW Asymmetrically tailed plasmid primer; adaptor; ds.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 1..4
FT FT /*tag= a
FT FT /label= Sticky end
FT FT 15..56
FT FT /*tag= b
FT FT /label= Sticky end
PN US5270185-A.
PD 14-DEC-1993.
PE 21-APR-1989; 341523.
PR 21-APR-1989; US-341523.
PR 12-AUG-1992; US-928855.
PA (HOFF.) HOFFMANN LA ROCHE INC.
PI Margolskee RF;
PT WPI: 93-404926/50.
PT Asymmetrically tailed plasmid primers - comprises cut,
PT double-stranded DNA plasmid, useful for producing cloning
PT libraries
PS Claim 5; Figure 3B, 13pp. English.
CC The inventors claim an asymmetrically tailed plasmid primer
CC comprising a cut, ds DNA plasmid contg. a functional origin of
CC replication and at least one functional selection marker gene,
CC termini of which contains an oligo (dT) extension, the other 3'
CC terminus of which contains an oligo (dC) or oligo (dG) extension
CC terminated by a 3' phosphate group. More specifically, the
CC plasmid primer is produced by EcoRI and HindIII cleavage
CC of plasmid pUC19, and each terminus produced by the cleavage
CC is joined to the complementary adaptor, which is 052733 or 052734.
QQ Sequence 56 BP; 1 A; 5 C; 8 G; 42 T;

```

Query Match	2.8%	Score 49;	DB 1;	Length 56;
Best Local Similarity	100.0%	Pred. No. 5.8e-09;		
Matches 49;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

Db 56 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCC 8

RESULT	24
T84975	
ID	T84975 standard; cDNA; 2378 BP.
AC	T84975;

DT	27-APR-1998	(first entry)	
DE	Rat glial cell derived neurotrophic factor receptor alpha cDNA.		
KW	Glial cell derived neurotrophic factor receptor alpha; GDNFR alpha;		
KW	GDNF; rat; kidney disease; glomerulonephritis; therapy; ds.		
OS	Rattus sp.		
FM	Key	location/Qualifiers	
FT	CDS	117..153	
FT		/*tag- a	
FT	sig_peptide	117..188	
FT		/*tag- b	
FT	mat_peptide	189..1550	
FI		/*tag- c	
PM	MO9733912-A2.		
PD	18-SEP-1997.		
PF	13-MAR-1997; 004363.		
PR	14-MAR-1996; US-618236.		
PR	14-MAR-1996; US-615902.		
PA	(GETH) GENENTECH INC.		
PI	Klein RD, Moore MM, Rosenthal A, Ryan AM;		
DR	WPI: 97-470819/43.		
DR	P-PSDS: W27327.		
PT	Isolated glial cell derived neurotrophic factor receptor alpha -		
PT	useful to develop products to diagnose and treat associated		
PT	disorders, particularly enteric nervous system or kidney disorders		
PS	Claim 23: Page 74-77, 100pp. English.		
CC	This cDNA codes for full-length rat glial cell derived neurotrophic		
CC	factor receptor alpha (GDNFR alpha) (see W27327), a novel		
CC	GPI-linked protein that is a ligand-binding component of the		
CC	receptor system for GDNF. It was isolated by expression cloning.		
CC	A cDNA library obtained from ventral midbrain tissue of E14 rat		
CC	embryos was generated in a cytomalovirus-based vector. cDNA		
CC	clones were transfected into COS 7 cells and expression of		
CC	putative GDNF receptors was detected by binding of iodinated GDNF.		
CC	A single positive pool was obtained, from which the cDNA clone was		
CC	isolated. An expression vector containing the cDNA can be used to		
CC	produce GDNFR alpha in transformed host cells. The invention		
CC	relates to novel uses of GDNF and its receptor. In particular, it		
CC	relates to native rat GDNFR alpha (see W27327), its variants and		
CC	soluble derivatives (extracellular domain), chimeric GDNFR alpha		
CC	and antibodies which bind to the GDNFR alpha, including agonist		
CC	and neutralising antibodies, as well as various uses for these		
CC	molecules. It also relates to assay systems for detecting ligands		
CC	to GDNFR alpha, systems for studying the physiological role of		
CC	GDNF, diagnostic techniques for identifying GDNF-related conditions		
CC	methods for identifying molecules homologous to GDNFR alpha, and		
CC	therapeutic techniques (claimed) for the treatment of GDNF-related		
CC	CC and GDNFR alpha-related conditions, particularly kidney disease		
CC	associated with glomerulonephritis and enteric nervous system		
CC	related disorders. Transgenic and knockout animals are also		
CC	claimed.		
Sequence	2378 BP; 617 A; 624 C; 598 G; 539 T;		

```

Query Match      2.8%; Score 49; DB 1; Length 2378;
Best Local Similarity 100.0%; Pred. NO. 2.7e-09;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

QY	1665	ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	1713
Dd	2250	ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2298

Db 2250 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2298

RESULT	25
RD	V33197 standard; cDNA; 3443 BP.
AC	V33197.
DT	21-DEC-1998 (first entry)
DE	Secreted protein DN747-7 cDNA.
KW	Secreted protein; DN747-7; human; ds
OS	Homo sapiens.
FT	Location/Qualifiers
FT	1599..2546
FT	/*tag= a

PN MO5838209-A2.
 PD 03-SEP-1998.
 PF 25-FEB-1998; U03697.
 PR 24-FEB-1998; US-028724.
 PR 26-FEB-1997; US-805819.
 PA (GENY) GENETICS INST INC.
 PI Agostino M, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
 PI Racie JA, Spaulding V, Treacy M;
 PI WPI: 98-481139/41.
 DR P-PSDB: W70325.
 PI New isolated polynucleotide(s) and encoded polypeptide(s) -
 PI obtained from human foetal kidney, adult colon, adult brain, foetal
 PI brain and placenta cDNA libraries.
 PS Claim 29: Page 75-77: 103pp: English.
 CC This full-length cDNA clone, designated DN747_7, codes for a novel
 CC human secreted protein (see W70325). It was isolated from a human
 CC foetal brain cDNA library using a method selective for cDNAs
 CC encoding secreted proteins, or was identified as encoding a
 CC secreted or transmembrane protein on the basis of computer analysis
 CC of the amino acid sequence of the encoded protein. Database
 CC searching revealed some similarity between DN747_7 and some known
 CC sequences. The invention provides new isolated polynucleotides
 CC (see V33190-99), from human foetal kidney, adult colon, adult brain,
 CC foetal brain and placenta cDNA libraries, that code for secreted
 CC proteins (see W70319-27). They can be isolated from composite
 CC deposit ATCC 98337 using specific probe sequences (see V33200-08).
 CC The clones can be used for recombinant production of the
 CC polypeptides, which may have activities such as e.g. nutritional
 CC activity, cytokine and cell proliferation/differentiation activity,
 CC immunostimulant or immunosuppressive, haematopoiesis regulating
 CC activity, tissue growth activity, actinin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, antiinflammatory activity,
 CC cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity, or other activities.
 SQ Sequence 3443 BP; 875 A; 934 C; 834 G; 766 T;

Query Match 2.8%; Score 49; DB 1; Length 3443;
 Best Local Similarity 100.0%; Pred. No. 2.5e-09;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1673 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGC 1721
 DB 1508 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGC 1556

RESULT 26
 VS4124
 ID VS4124 standard; DNA; 2519 BP.
 AC VS4124;
 DT 17-DEC-1998 (first entry)
 DE Human membrane protein BA0306 coding sequence.
 KW Membrane protein; BA0306; arteriosclerosis; coronary restenosis;
 KW therapy: human; ds.
 OS Homo sapiens.
 FH Key
 PR Location/Qualifiers
 FT 1..1785
 FT /tag= a
 FT /transl_except= (pos: 289..291, aa: Xaa)
 FT /transl_except= (pos: 1015..1017, aa: Xaa)
 FT /transl_except= (pos: 1492..1494, aa: Xaa)
 FT /transl_except= (pos: 1540..1542, aa: Xaa)
 FT /transl_except= (pos: 1582..1584, aa: Xaa)
 FT /transl_except= (pos: 1609..1611, aa: Xaa)
 FT /transl_except= (pos: 1693..1695, aa: Xaa)
 FT /note= "Xaa= unspecified amino acid"

MO9838305-A1.
 PD 03-SEP-1998.
 PF 27-FEB-1998; J00835.
 PR 25-FEB-1998; JP-062263.
 PR 28-FEB-1997; JP-062259.
 PA (NISH.) JAPAN TOBACCO INC.

PI Nakamura Y, Tanaka T, Tsukada S;
 DR WPI: 98-481206/41.
 DR P-PSDB: W74580.
 PT Membrane protein(s) BA0306 and BA2303 - useful for, e.g. treatment
 PT and prevention of arteriosclerosis and restenosis
 PS Claim 16; Page 90-95; 141pp: Japanese.
 CC This sequence encodes the human BA0306 membrane protein of the invention.
 CC The invention also relates to the human BA2303 membrane protein. The two
 CC membrane proteins are specifically expressed in mammals during
 CC arteriosclerosis and coronary restenosis. The membrane proteins,
 CC fragments of them, and antibodies against them are useful in the
 CC treatment and prevention of arteriosclerosis and restenosis. Transgenic
 CC mice expressing the extracellular region of the membrane proteins are
 CC useful as models for studying these disorders.
 SQ Sequence 2519 BP; 766 A; 439 C; 492 G; 809 T;

Query Match 2.8%; Score 49; DB 1; Length 2519;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1673 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGC 1721
 DB 2434 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGC 2482

RESULT 27
 X52242
 ID X52242 standard; DNA; 1770 BP.
 AC X52242;
 DT 25-JUN-1999 (first entry)
 DE Protein PRO33 cDNA clone DNA34436-1238.
 KW Secreted protein; transmembrane protein; human; enterocolitis;
 KW Zollinger-Ellison syndrome; gastrointestinal ulceration;
 KW congenital microvillus atrophy; skin disease; cell growth;
 KW abnormal keratinocyte differentiation; psoriasis; epithelial cancer;
 KW Parkinson's disease; Alzheimer's disease; ALS; neuropathy;
 KW fibromodulin; dermal scarring; Usher Syndrome; Atrophia areata;
 KW anti-thrombotic; wound healing; tissue repair; ss.
 OS Homo sapiens.
 PN MO9914328-A2.
 PD 25-MAR-1999.
 PF 16-SEP-1998; U19330.
 PR 25-NOV-1997; US-056840.
 PR 17-SEP-1997; US-059113.
 PR 17-SEP-1997; US-059115.
 PR 17-SEP-1997; US-059117.
 PR 17-SEP-1997; US-059119.
 PR 17-SEP-1997; US-059121.
 PR 17-SEP-1997; US-059122.
 PR 17-SEP-1997; US-059184.
 PR 18-SEP-1997; US-059263.
 PR 18-SEP-1997; US-059266.
 PR 15-OCT-1997; US-062125.
 PR 17-OCT-1997; US-062285.
 PR 17-OCT-1997; US-062287.
 PR 21-OCT-1997; US-063486.
 PR 24-OCT-1997; US-062814.
 PR 24-OCT-1997; US-062816.
 PR 24-OCT-1997; US-063045.
 PR 24-OCT-1997; US-063120.
 PR 24-OCT-1997; US-063121.
 PR 24-OCT-1997; US-063127.
 PR 24-OCT-1997; US-063128.
 PR 27-OCT-1997; US-063329.
 PR 27-OCT-1997; US-063327.
 PR 28-OCT-1997; US-063541.
 PR 28-OCT-1997; US-063542.
 PR 28-OCT-1997; US-063544.
 PR 28-OCT-1997; US-063549.
 PR 28-OCT-1997; US-063550.
 PR 28-OCT-1997; US-063564.
 PR 29-OCT-1997; US-063435.

29-OCT-1997: US-063704.
PR 29-OCT-1997: US-063732.
PR 29-OCT-1997: US-063738.
PR 29-OCT-1997: US-063734.
PR 29-OCT-1997: US-064215.
PR 29-OCT-1997: US-063735.
PR 31-OCT-1997: US-063870.
PR 31-OCT-1997: US-064103.
PR 03-NOV-1997: US-064248.
PR 07-NOV-1997: US-064809.
PR 12-NOV-1997: US-065186.
PR 17-NOV-1997: US-065846.
PR 18-NOV-1997: US-065693.
PR 21-NOV-1997: US-065120.
PR 21-NOV-1997: US-066364.
PR 24-NOV-1997: US-066772.
PR 24-NOV-1997: US-066466.
PR 24-NOV-1997: US-066770.
PR 24-NOV-1997: US-066511.
PR 24-NOV-1997: US-066453.
PA (GETH) GENENTECH INC.
PI Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J:
DR WPI: 99-229533/19.
DR P-PSDB: Y13371.
PT New isolated human genes and polypeptides used in, e.g. treatment of
PT gastrointestinal ulceration
PS Claim 2: Fig 57: 320pp: English.
CC X52213-7: encode secreted and transmembrane human proteins, and are
CC obtained from cDNA libraries, prepared from fetal lung, fetal kidney,
CC fetal brain, fetal liver and fetal retina. The encoded polypeptides
CC have specific uses based on their homology to known polypeptides,
CC e.g. PRO211 and PRO217 can be used for disorders associated with the
CC preservation and maintenance of gastrointestinal mucosa and the repair
CC of acute and chronic mucosal lesions (e.g. enterocolitis,
CC Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital
CC microvillus atrophy), skin diseases associated with abnormal keratinocyte
CC differentiation (e.g. psoriasis, epithelial cancers such as lung squamous
CC cell carcinoma of the vulva and gliomas), potent effects on cell growth
CC and development, diseases related to growth or survival of nerve cells
CC including Parkinson's disease, Alzheimer's disease, ALS, neuropathies or
CC cancer. PRO265 can be used as for fibromodulin, e.g. for reducing dermal
CC scarring. PRO264 can be used as a target for anti-tumor drugs. PRO333
CC may be used in the treatment of Usher Syndrome or Atrophila areata;
CC PRO209 can be used as an anti-thrombotic agent; PRO287 polypeptides
CC and portions may have therapeutic applications in wound healing and
CC tissue repair; PRO317 can be used for treating problems of the kidney,
CC uterus, endometrium, blood vessels, or related tissue, e.g. in the
CC heart of genital tract. 621 A: 313 C: 329 G: 507 T:
SQ Sequence 1770 BP: 621 A: 313 C: 329 G: 507 T:

Query Match 2.8%; Score 49; DB 1; Length 1770;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1665 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
DB 1663 ACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1711

RESULT 28
O47967
ID O47967 standard: cDNA: 1586 BP.
AC O47967;
DT 25-MAR-1994 (first entry)
DE Rape acyl-ACP thioesterase (Clone pNL3) coding sequence.
KW Enzyme precursor: acyl-ACP thioesterase; plant; transformation;
KM fatty acid; yield; composition; seed oil; Brassica napus; ss.
OS Brassica napus.
FH Key Location/Qualifiers
FT 169..1269
FT CDS
FT /tag= a
FT /product= Enzyme precursor.

FT W09318158-A. /note= "Has acyl-ACP thioester activity."
PN 16-SEP-1993.
PD 03-MAR-1993; G00432.
PF 03-MAR-1992; GB-004583.
PR (UNIL) UNILEVER NV.
PA (UNIL) UNILEVER PLC.
PI Hellyer SA, Loader NM, Safford R, Slabas AR:
DR WPI: 93-303472/38.
DR P-PSDB: R41674.
PT cDNA encoding enzyme with acyl-ACP-thioesterase activity
PT useful for transforming plants, for altering properties e.g. seed
PT storage oil composition
PS Claim 2: Figure 2: 48pp: English.
CC The nucleotide sequence is useful for transforming plant cells and
CC altering that plants characteristics. Characteristics which may be
CC altered are (1) altering the level of acyl-ACP thioesterase
CC activity; (2) fatty acid yield and/or fatty acid composition and (3)
CC yield and/or composition of seed storage oil.
SQ Sequence 1586 BP: 491 A: 324 C: 339 G: 432 T:

Query Match 2.8%; Score 48; DB 1; Length 1586;
Best Local Similarity 100.0%; Pred. No. 5.6e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
DB 1521 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1568

RESULT 29
O58663
ID O58663 standard: cDNA to mRNA: 699 BP.
AC O58663;
DT 14-OCT-1994 (first entry)
DE Human superoxide dismutase gene.
KW Superoxide dismutase; SOD; transformation; production; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS
FT 68..329
FT /tag= a
FT /product= superoxide_dismutase
FT 625..699
FT /tag= b
PN poly-a-signal
PD J06046860-A.
PD 22-FEB-1994.
PF 23-SEP-1992; 279193.
PR 23-SEP-1992; JP-279193.
PA (NIPK) NIPPON KAYAKU KK.
DR WPI: 94-097023/12.
DR P-PSDB: R47062.
PT DNA encoding human superoxide dismutase (SOD) and transformed
PT microorganism - useful for prodn. of SOD
PS Claim 1: Page 7: 10pp: Japanese
CC The sequence encodes a superoxide dismutase. The gene may be used to
CC produce superoxide dismutase in increased yields by transforming it
CC into E.coli.
SQ Sequence 699 BP: 242 A: 127 C: 181 G: 149 T:

Query Match 2.8%; Score 48; DB 1; Length 699;
Best Local Similarity 100.0%; Pred. No. 7.8e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
DB 622 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 669

RESULT 30
O75361
ID O75361 standard: cDNA to mRNA: 1490 BP.

```

AC 075361:
DT 01-JUL-1995 (first entry)
DE Gibberellin-20-oxidase.
KW Gibberellin-20-oxidase; plasmid PAT253; transgenic plant;
  crop improvement; ss.
OS Arabidopsis thaliana.
FH key Location/Qualifiers
FT cds 98..1231
FT /*tag= a
FT /*note= "GA-20-oxidase"
PN MO9428141-A.
PD 08-DEC-1994.
PF 24-MAY-1994; E01664.
PR 28-MAY-1993; GB-011147.
PA (LONG-) LONG ASHTON RES STATION.
PI Graebe JE, Hedden P, Lange T, Phillips A;
  WP1: 95-022813/03.
DR P-PSDB: R64144.
PT New DNA encoding gibberellin-20-oxidase and related vectors -
  also host cells and transgenic plants, used to modulate plant
  growth and development, e.g. anti-sense constructs will reduce
  vegetative growth.
PS Claim 10: Page 78-79; 97pp: English.
CC Clone PAT353 containing DNA encoding gibberellin-20-oxidase may be
  used to construct chimeric genes and binary vectors to transform
  plants by usual methods, e.g. Agrobacterium-mediated transfer.
CC propagated transgenic plants show modulated growth and development
  e.g. reducing GA-20-oxidase expression with antisense constructs
  reduces vegetative growth in cereals and grasses, improving
  resistance to frost, etc. Sense constructs improve fruit set and
  growth, extended stems and leaves, induce flowering and overcome
  dormancy, etc. Preferred transgenic plants are maize and wheat.
CC Sequence 1490 BP; 486 A; 330 C; 274 G; 400 T;
  SO
Query Match 2.8%; Score 48; DB 1; Length 1490;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1666 CTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
  |||||||
DB 1421 CTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1468
RESULT 31
ID 099552 standard; cDNA to mRNA; 861 BP.
AC 099552:
DT 29-APR-1996 (first entry)
DE Thrombopoietin coding sequence.
KW Thrombopoietin; TPO; humoral factor; platelet; antibody; therapy; AIDS;
  thrombocytopenia; hypoplastic anaemia; thrombotic thrombocytopenia;
  disseminated intravascular coagulation syndrome; ds.
OS Homo sapiens.
FH key Location/Qualifiers
FT cds 25..783
FT /*tag= a
FT /*product= thrombopoietin
FT signal_peptide 25..87
FT /*tag= b
FT mat_peptide 88..783
FT /*tag= c
PN MO9521919-A2.
PD 17-AUG-1995.
PF 14-FEB-1995; J00208.
PR 14-FEB-1994; JP-039090.
PR 14-MAR-1994; US-212164.
PR 25-MAR-1994; JP-079842.
PR 01-APR-1994; US-221020.
PR 01-JUN-1994; JP-155126.
PR 15-JUN-1994; JP-167328.
PR 20-JUL-1994; US-278083.
PR 17-AUG-1994; JP-193169.

```

```

PR 17-AUG-1994; JP-227159.
PR 18-AUG-1994; JP-193916.
PR 11-OCT-1994; US-320300.
PR 01-NOV-1994; JP-304167.
PR 01-DEC-1994; JP-298669.
PR 22-DEC-1994; US-361811.
PR 28-DEC-1994; JP-341200.
PR 31-JAN-1995; US-381478.
PA (KIRI) KIRIN BREWERY KK.
PI Akahori H, Iwamatsu A, Kato T, Kuroki R, Miyazaki H;
  Muto T, Ohgami K, Shimizu T;
  WP1: 95-283120/38.
DR P-PSDB: R81378.
PT Thrombopoietin polypeptide which specifically stimulates or
  increases platelet production - useful in treatment of
  thrombocytopenia, also related DNA and vectors
PS Claim 19; Page 337-338; 383pp: English.
CC This sequence represents the coding sequence for a thrombopoietin (TPO)
  protein. TPO is a humoral factor capable of promoting platelet
  production. The encoded protein can be purified by using a gel affinity
  column where Mpl has been coupled to a resin. 099551 and 099553 also
  encode TPO proteins. The DNA sequences are inserted into vectors which
  are used to transform prokaryotic and eukaryotic host cells. Using the
  proteins encoded by these sequences, and derivatives of them, antibodies
  specifically immunoreactive with a TPO protein can be created. The
  antibodies, DNA sequences and vectors are used to isolate the protein
  sequences. The TPO proteins can then be used in the treatment of
  platelet disorders. These include thrombocytopenia, hypoplastic anaemia,
  AIDS, disseminated intravascular coagulation syndrome and thrombotic
  thrombocytopenia.
CC Sequence 861 BP; 243 A; 257 C; 197 G; 164 T;
  SO
Query Match 2.8%; Score 48; DB 1; Length 861;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1666 CTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
  |||||||
DB 783 CTTAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAA 830
RESULT 32
ID T90174 standard; cDNA; 1474 BP.
AC T90174:
DT 20-APR-1998 (first entry)
DE Oil seed rape cysteine protease cDNA clone CDCY566.
KW Promoter; oil seed rape; cDNA clone CDCY566;
  cysteine protease; disrupter protein; plant; tolerance; herbicide;
  insect pest; fungal disease; improved yield; improved quality;
  fertility control; ss.
OS Brassica napus.
FH key Location/Qualifiers
FT cds 33..1157
FT /*tag= a
PN MO9735983-A2.
PD 02-OCT-1997.
PF 18-MAR-1997; G00729.
PR 22-MAR-1996; GB-006062.
PA (ZENE) ZENECA LTD.
PI Greenland AT, Jepson I, Thomas DRP;
  DR WP1: 97-489646/45.
DR P-PSDB: W27446.
PT Cysteine protease promoter - obtained from oil seed rape, useful to
  develop plants with improved agronomic characteristics
PS Claim 23; Fig 17; 137pp: English.
CC The present sequence encodes an oil seed rape cysteine protease,
  the promoter for which can restrict the expression of a disrupter
  protein gene to a suitable stage of plant development, to provide
  plants with novel agronomic features, e.g. tolerance to herbicides,
  insect pests and fungal diseases, improved yield and/or quality of
  harvested product, and novel mechanisms for the control of plant

```


CC fertility: 1474 BP; 528 A; 265 C; 316 G; 365 T;
 SD Sequence

Query Match 2.8%; Score 48; DB 1; Length 1474;
 Best Local Similarity 100.0%; Pred. No. 6.7e-09;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1566 CTTAAAAA
 DB 1373 CTTAAAAA

RESULT 33

V21646
 ID V21646 standard; cDNA; 2312 BP.

AC V21646;
 DT 21-JUL-1998 (first entry)
 DE T. gondii chorismate synthase encoding cDNA.
 KW Chorismate synthase; enzyme; Apicomplexan parasite; diagnosis;
 KM metabolic pathway; parasitic disease; psda gene; toxoplasma gondii;
 OS Toxoplasma gondii.
 FH Key Location/Qualifiers
 FT CDS 152..1172
 FT /tag=a
 FT /product="chorismate synthase enzyme"

PT W09803661.A2.
 PD 29-JAN-1998.
 PF 18-JUL-1997; U12497.
 PR 13-JUN-1997; US-049620.
 PR 19-JUL-1996; US-022309.
 PR 23-DEC-1996; US-773302.
 PR 17-MAR-1997; US-040849.
 PA (ARCH-) ARCH DEV CORP.
 PI Johnson JJ, McLeod RLM, Mels L, Roberts CW, Roberts F;
 DR WPI: 98-120782/11.
 DR P-PSDB: W53105.

PT Use of Apicomplexan parasite components - from plant-like metabolic
 PT pathways for developing products to treat, diagnose or prevent
 PT parasitic diseases
 PS Claim 12: Pages 172-175, 211pp; English.
 CC This cDNA encodes a T. gondii chorismate synthase enzyme. A component of
 CC a plant-like metabolic pathway in an Apicomplexan parasite used for
 CC producing a composition that interferes with the growth or survival of
 CC the Apicomplexan parasite is selected from this 2312 base pair T. gondii
 CC chorismate synthase enzyme sequence or a 1837 base pair P. falciparum
 CC chorismate synthase enzyme sequence. The pathway does not involve the
 CC psda gene or PPI phosphotransferase and the component is not encoded by
 CC the plasmid genome, and is not generally operative in animals. The
 CC component comprises a transit peptide when the pathway involves transport
 CC of proteins into or out of an organelle. An Apicomplexan parasite
 CC comprising a gene encoding a component of a metabolic pathway which
 CC operates for the growth and survival of the parasite, where the gene is
 CC altered can be used to construct a vaccine for protecting an animal
 CC against infection by the parasite. The invention provides methods for
 CC identifying such components and for developing antibodies and inhibitors
 CC to the component. The products and methods can be used to develop
 CC products to treat, diagnose and prevent diseases due to Apicomplexan
 CC parasites such as Toxoplasma gondii causing toxoplasmosis, Plasmodia
 CC causing malaria, Cryptosporidia causing cryptosporidiosis, Eimeria
 CC causing eimeriosis, Babesia causing babesiosis, Theileria causing
 CC theileriosis and Neospora caninum.
 SQ Sequence 2312 BP; 548 A; 640 C; 624 G; 500 T;

Query Match 2.8%; Score 48; DB 1; Length 2312;
 Best Local Similarity 100.0%; Pred. No. 6.1e-09;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1566 CTTAAAAA
 DB 2249 CTTAAAAA

RESULT 34

ID V18187 standard; cDNA to mRNA; 5503 BP.

AC V18187;
 DT 28-AUG-1998 (first entry)
 DE Fanconi anaemia of complementation group A gene.
 KW Fanconi anaemia of complementation group A; FA-A; genetic defect;
 KM prenatal FA-A; FA-A carrier detection; disease diagnosis; ss.
 OS Homo sapiens.

FH Key Location/Qualifiers
 FT CDS 32..4399
 FT /tag=a
 FT /product= FA-A

PN W09814462-A1.
 PD 09-APR-1998.
 PF 03-OCT-1997; U18010.
 PR 04-OCT-1996; US-726012.
 PA (FANC-) FANCONI ANEMIA RES FUND INC.
 PI Joenje H, Lo Ten Foe JR;
 DR WPI: 98-240012/21.
 DR P-PSDB: W48663.
 PT DNA for Fanconi Anaemia complementation group A - useful for e.g.
 PT developing products for diagnosis and screening of disease and gene
 PT therapy
 PS Claim 4: Fig 2; 63pp; English.
 CC This sequence encodes the Fanconi anaemia of complementation group A
 CC (FA-A) protein of the invention. The DNA's may be used to complement a
 CC genetic defect in a cell (especially the FA-A gene). The products can be
 CC used for screening (especially prenatal FA-A), detection of FA-A carriers
 CC and FA-A disease diagnosis
 SQ Sequence 5503 BP; 1208 A; 1527 C; 1492 G; 1276 T;

Query Match 2.8%; Score 48; DB 1; Length 5503;
 Best Local Similarity 100.0%; Pred. No. 5.1e-09;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1566 CTTAAAAA
 DB 5446 CTTAAAAA

RESULT 35

V37789/C
 ID V37789 standard; DNA; 90 BP.

AC V37789;
 DT 09-SEP-1998 (first entry)
 DE Analytical solid phase nucleic acid detection hybridisation probe 16.
 KW Analytical solid phase: detection; hybridisation; probe; target;
 KW diagnosis; screening; disease; drug sensitivity; transplantation organ;
 KW food industry; food poisoning; ss.
 OS Synthetic.
 PN W09811210-A1.
 PD 19-MAR-1998.
 PF 12-SEP-1997; J03232.
 PR 13-SEP-1996; JP-243720

PA (MOLE-) LAB MOLECULAR BIOPHOTONICS.
 PI Abe S, Sato Y;
 DR WPI: 98-271664/24.
 DR Analytical solid phase for detecting nucleic acids - contains a base
 PT sequence which hybridises with polynucleotide sequence of the
 PT target, and a set of probes immobilised on the solid phase via a
 PT linker

PS Example 7: Page 21-22; 37pp; Japanese.
 CC The present sequence represents a probe used in an example from the
 CC present invention. The present invention describes an analytical solid
 CC phase method for detecting nucleic acids. The method comprises obtaining
 CC a base sequence which hybridises with the polynucleotide sequence of the
 CC target, and a set of probes immobilised on the solid phase via a linker
 CC which is enzymatically ligated during hybridisation. The product is used
 CC for the detection of specific nucleic acids. Possible fields of


```
FT      sig_peptide      /*tag= a
FT      563.601
FT      /*tag= b
FT      mat_peptide      /tag= c
FT      602.1111
PN      WO980404-A2.
PD      17-SEP-1998.
PF      09-MAR-1998: U04601.
PR      06-MAR-1998: US-016321.
PR      11-MAR-1997: US-815381.
PA      (GEMT ) GENETICS INST INC.
PI      Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
PI      Racine LA, Spaulding V, Treacy M.
PI      WPI: 98-520802/44.
DR      P-PSDB: W69423.
PT      New isolated polynucleotides and secreted proteins - useful as, e.g.
PT      nutritional additives, immunostimulators, haemotopoiesis regulators
PS      Claim 1: Page 56: 11pp: English.
CC      This full-length cDNA clone, designated ax318.3, codes for a novel
CC      human secreted protein (see W69423), i.e. ax318.3 protein. It was
CC      isolated from a human adult testis cDNA library using methods which
CC      are selective for cDNAs encoding secreted proteins, or was
CC      identified as encoding a secreted or transmembrane protein on the
CC      basis of computer analysis of the encoding protein. Homology is
CC      shown to some database sequences. The invention provides isolated
CC      polynucleotides (see V58754-63) obtained from human adult testis,
CC      brain, retina or placenta, or from foetal kidney or brain cDNA
CC      libraries. These are all deposited as ATCC 98353. They encode
CC      novel human secreted proteins (see W69423-33) that may have e.g.
CC      nutritional activity, immune stimulating or suppressing activity,
CC      haematopoiesis regulating activity, tissue growth activity,
CC      activin/inhibin activity, chemotactic/chemokinetic activity,
CC      haemostatic and thrombolytic activity, receptor/ligand activity,
CC      antiinflammatory activity, cadherin/tumour invasion suppressor
CC      activity, tumour inhibition or other activities. They may also be
CC      used for diagnostic purposes.
SQ      Sequence 1328 BP: 446 A: 307 C: 282 G: 293 T:
```

```
Query Match      2.8%; Score 48; DB 1: Length 1328;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 48: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
```

```
OY      1669 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCG 1716
DB      452 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCG 499
```

```
RESULT 39
V52969 ID V52969 standard: cDNA: 1373 BP.
AC V52969:
DT 11-JAN-1999 (first entry)
DE Human G1 protein isoform beta (CASH-beta) cDNA.
KW G1 protein; CASH-beta; human; caspase homologue; Fas receptor;
KW modulator; apoptosis; cell death; inflammation; tumour; HIV;
KW therapy; ss.
OS Homo sapiens.
FH key Location/Qualifiers
FT CDS 482..1137
FT /*tag= a
PN WO9839435-A1.
PD 11-SEP-1998.
PF 26-FEB-1998: IL0098.
PR 01-MAY-1997: IL-120759.
PR 03-MAR-1997: IL-120367.
PA (YEDA ) YEDA RES & DEV CO LTD.
PI Brodalski V, Goltsev Y, Kovalenko A, Varfolomeev E,
PI Wallach D.
PI WPI: 98-495842/42.
DR P-PSDB: W78904.
PT New DNA encoding isoforms of G1 protein which bind MORT-1 - and
```

```
PT regulate the effects of FAS and tumour necrosis factor receptors,
PT useful for killing of cells e.g. HIV and cancer cells
PS Claim 4: Fig 2: 13pp: English.
CC This cDNA sequence codes for the beta isoform (see W78904) of novel
CC human G1 protein. It was isolated from a human skin fibroblast
CC cDNA library using a cDNA probe corresponding to the G1 sequence.
CC G1-beta (also called CASH beta, CASH being caspase homologue)
CC and a longer isoform, G1-alpha (see W78903), represent 2 splice
CC variants of the G1 protein. These G1 proteins are capable of
CC binding to, or interacting directly or indirectly, with MORT-1
CC or with MORT-binding proteins such as Mch4 (CASP-10) and MCH
CC (CASP-8), and thereby of binding to the intracellular domain of
CC the FAS-R receptor, to which MORT-1 binds, or of binding to the
CC intracellular domain of the p55 tumour necrosis factor (TNF)
CC receptor, to which TRADD binds and to which TRADD protein MORT-
CC 1 binds. Hence, they are considered as mediators or modulators of
CC FAS-R having a role in e.g. the signalling process that is initiated
CC by the binding of FAS ligand to FAS-R, and also having a role in
CC the signalling process initiated by the binding of TNF to p55-R.
CC G1 DNA (I) and polypeptide (II), vectors and fragments are used to
CC regulate cell death or inflammatory processes. (II) is used to
CC inhibit cell death, and its inhibitors augment/enhance the
CC processes. (I) and (II) regulate the FAS-R ligand or TNF effect on
CC cells carrying an FAS-R or p55-R. Tumour, HIV-infected or other
CC diseased cells can be treated using a viral vector encoding a viral
CC surface protein, which binds a specific cell surface receptor and a
CC sequence encoding (II), which kills the cell. Antisense
CC oligonucleotides, introduced using the above vector, block the
CC expression of (II) and can also regulate the above effects. These
CC effects can also be regulated using a vector encoding a lysozyme
CC that interacts with a cellular mRNA encoding (II), and allows (II)
CC expression.
SQ Sequence 1373 BP: 417 A: 283 C: 335 G: 338 T:
```

```
Query Match      2.8%; Score 48; DB 1: Length 1373;
Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 48: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
```

```
OY      1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
DB      1310 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1357
```

```
RESULT 40
V64590 ID V64590 standard: DNA: 1579 BP.
AC V64590:
DT 29-JAN-1999 (first entry)
DE Tobacco chili DNA.
KW ChlD: magnesium chelatase subunit D; tobacco; transgenic plant; activity;
KW subunit I; subunit H; effector; detectable; ss.
OS Nicotiana tabacum.
PN DE19717656-A1.
PD 29-OCT-1998.
PF 25-APR-1997: 017656.
PR 25-APR-1997: DE-017656.
PA (AGRE ) HOECHST-SCHERING AGREVO GMBH.
PI Graefe S, Grimm B, Haendel F, Papenbrock J, Schmidt F,
PI Streiber W.
PI WPI: 98-569720/49.
PT New nucleic acid encoding plant magnesium chelatase subunit D
PT protein - used to amplify nucleic acid, produce antibodies and
PT transgenic plants and to identify plant magnesium chelatase
PT effectors
PS Example 6, Page 23-24: 30pp: German.
CC This sequence encodes a tobacco magnesium chelatase subunit I (chl1)
CC which is used in a method to amplify, isolate or identify a nucleic
CC acid molecule coding for a chlD protein or a biologically active
CC fragment, to produce antibodies and to produce transgenic plants. The
CC chlD protein is used for magnesium chelatase activity determination.
CC Determination of the interaction of plant magnesium chelatase subunits
CC comprises transforming a host cell with the nucleic acid and with at
```

CC least one DNA sequence encoding another magnesium chelatase subunit, in
CC such a way that the interaction of the magnesium chelatase subunits
CC produces a detectable signal, preferably by activation of a reporter
CC gene. Determination of the activity of a plant magnesium chelatase
CC comprises contacting the CHLD protein with the gene products of DNA
CC sequences coding for magnesium chelatase subunits I and H in such a
CC way that the enzymatic activity of the magnesium chelatase subunits
CC produces a detectable signal. The above methods are used to identify
CC plant magnesium chelatase effectors.
Sequence 1579 BP; 490 A; 285 C; 368 G; 436 T;

Query Match 2.8%; Score 48; DB 1; Length 1579;
Best Local Similarity 100.0%; Pred. No. 6.6e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
|||||
DB 1508 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1555

RESULT 41

X33812
ID X33812 standard; DNA; 2120 BP.
AC X33812.
DT 25-JUN-1999 (first entry)
DE Coding sequence for human secreted protein cb96.10.
KW Secreted protein; human; nutritional activity; cytokine; tissue growth;
KW cell proliferation; cell differentiation; immune stimulant; chemotaxis;
KW immune suppressant; haematopoiesis regulator; activin; inhibin; cadherin;
KW chemokinesis; haemostasis; thrombolysis; anti-inflammatory; gene therapy;
KW tumour invasion suppressor; tumour inhibitor; ss.
OS Homo sapiens.
PN WO9913066-A1.
PD 13-MAR-1999.
PF 03-SEP-1998; U18724.
PR 03-SEP-1997; US-929007.
PA (GENY) GENETICS INST INC.
PI Agostino MJ, Evans C, Jacobs K, Lavalie ER, McCoy JM,
PI Merberg D, Racie LA, Spaulding V, Treacy M;
PI WPI: 98-229235/19.
DR P-PSDB: Y05319.

PT New polynucleotides encoding secreted human proteins
PS Claim 17; Page 80-81; 96pp; English.
CC This sequence encodes a human secreted protein of the invention. The
CC secreted proteins were obtained from human adult placenta, foetal brain,
CC adult testes or adult blood cDNA libraries. The polynucleotides (PNS) and
CC proteins are predicted to have biological activities which would make
CC them suitable for treating, preventing or ameliorating medical conditions
CC in humans and animals, although no supporting data is given. Suggested
CC activities include nutritional activity, cytokine and cell
CC proliferation/differentiation activity, immune stimulating
CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC invasion suppressor activity, and tumour inhibition activity. The PNS are
CC also stated to be useful for gene therapy. The PNS are
Sequence 2120 BP; 667 A; 347 C; 403 G; 703 T;

Query Match 2.8%; Score 48; DB 1; Length 2120;
Best Local Similarity 100.0%; Pred. No. 6.2e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
|||||
DB 2042 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2089

RESULT 42

X37521
ID X37521 standard; cDNA; 1090 BP.

AC X37521;
DT 06-JUL-1999 (first entry)
DE Human secreted protein cDNA fragment containing gene 71.
KW Human; secreted protein; treatment; prevention; protein therapy; AIDS;
KW gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;
KW developmental abnormality; fetal deficiency; blood disorder; leukemia;
KW immune system disease; autoimmune disease; hepatic disease; lymphoma;
KW renal disease; inflammation; allergy; Alzheimer's disease; schizophrenia;
KW cognitive disorder; prostate disease; skeletal; cardiac; muscle disorder;
KW pulmonary disorder; transplant rejection; osteoclast; osteoporosis;
KW arthritis; malignancy; digestive; endocrine; infection; ss.
OS Homo sapiens.
PN WO9918208-A1.
PD 15-APR-1999.
PF 01-OCT-1998; U20775.
PR 02-OCT-1997; US-060884.
PR 02-OCT-1997; US-060833.
PR 02-OCT-1997; US-060836.
PR 02-OCT-1997; US-060837.
PR 02-OCT-1997; US-060838.
PR 02-OCT-1997; US-060839.
PR 02-OCT-1997; US-060843.
PR 02-OCT-1997; US-060862.
PR 02-OCT-1997; US-060866.
PR 02-OCT-1997; US-060874.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Carter KC, Duan DR, Endress GA, Peng P, Ferrie AM,
PI Florence KA, Greene JM, Janat F, Lafleur DW, Ni J,
PI Rosen CA, Ruben SM, Shi Y, Young P, Yu G;
PI WPI: 99-264022/22.
DR P-PSDB: Y07922.

PT New isolated human genes and the secreted polypeptides they encode
PS Claim 1a; Page 263; 368pp; English.
CC This invention describes novel isolated human genes and the secreted
CC proteins they encode. The products of the invention are useful for
CC preventing, treating or ameliorating medical conditions, e.g. by protein
CC or gene therapy. Also pathological conditions can be diagnosed by
CC determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 101 polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, tumours,
CC neurodegenerative disorders, developmental abnormalities and fetal
CC deficiencies, blood disorders, leukemias, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, lymphomas, inflammation,
CC allergies, Alzheimer's and cognitive disorders, schizophrenia, prostate
CC disease, skeletal or cardiac muscle disorders, pulmonary disorders,
CC transplant rejection, disorders involving osteoclasts such as
CC osteoporosis, arthritis or malignancies, digestive/endocrine disorders,
CC infections and AIDS. The human secreted proteins of the invention are
CC represented in X37451-X37552.
Sequence 1090 BP; 256 A; 332 C; 295 G; 194 T;

Query Match 2.8%; Score 48; DB 1; Length 1090;
Best Local Similarity 100.0%; Pred. No. 7.1e-09;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1666 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
|||||
DB 1033 CTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1080

RESULT 43

N91576
ID N91576 standard; cDNA; 1239 BP.
AC N91576;
DT 28-FEB-1990 (first entry)
DE Sequence of clone 12R.
KW cDNA clone; clone 12R; Lol PI; Poaceae; Gramineae; Lolium perenne.
OS Lolium perenne.
PN WO8909260-A.

```

OS      synthetic.
FH      Key
FH      poly_a-signal
FT      1671..1676
FT      /tag=
FT      a
FT      165..1006
FT      /tag=
FT      b
PI      J02111796-A.
PI      24-APR-1990.
PR      21-OCT-1988; 264097.
PR      21-OCT-1988; JP-264097.
PR      (TOFU) Toa Nenryo Kogyo KK.
PR      MPI: 90-169110/22.
DR      P-PSDB; R05084.
PI      Amino acid hydrolase activating protein -
PI      isolated from mammal brain and activates tyrosine hydrolase
PI      and tryptophan hydrolase
PS      Disclosure; P: Japanese.
CC      Probable error in specification at posn 168. C residue should read
CC      G (to give Met codon Arg instead of ATC)
CC      Encodes a protein characterised by Mr of 27KD (by SDS-PAGE), iso-
CC      electric point of 4.6 and an N-terminal amino acid sequence beginning
CC      MetGlyAsp-.
CC      Protein activates Tyr hydrolase and Trp hydrolase in the presence
CC      of calcium ions and calmodulin-dependent protein phosphatase type II.
CC      Useful as a pharmaceutical analysis reagent.
CC      See also Q04591 and Q04692.
SO      Sequence 1834 BP; 546 A; 454 C; 471 G; 363 T;

Query Match      2.7%; Score 47; DB 1; Length 1834;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1667 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
Db 1695 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1741
|||||
|||||

RESULT 46
N80489
ID      N80489 standard; DNA: 372 BP.
AC      N80489.
DT      16-NOV-1990 (first entry)
DE      Cowpea trypsin inhibitor C-terminal coding sequence of PAgCl.
KW      Bowman-Birk cowpea trypsin inhibitor (CPTI); Insect resistance;
KW      PAgCl; ss.
OS      Vigna unguiculata.
PI      EP-272144-A.
PI      22-JUN-1988.
PI      18-DEC-1987; 311201.
PI      19-DEC-1986; GB-030448.
PI      02-DEC-1987; GB-025610.
PI      (AGRI-) Agric Genetics Ltd.
PI      Hilder VA, Gatehouse AMR, Boulter D;
PI      MPI: 88-169408/25.
PT      Recombinant DNA for transfer of specific genes into plants -
PT      esp for imparting insect resistance to plants and crops.
PS      Disclosure; P: English.
CC      PAgCl encodes the C-terminal half of a CPTI. The entire protein
CC      is required for insect resistance and so this sequence was used as
CC      a probe for in situ colony hybridisation. Full-length coding
CC      sequences isolated in this way are used to transfer insect
CC      resistance into economically important crops such as maize, rice,
CC      soybeans, vines, sugar beet, cotton, fruit and vegetables.
CC      See also P80887-P80889, N80490-N80491 and N82300.
SO      Sequence 372 BP; 148 A; 58 C; 70 G; 96 T;

Query Match      2.7%; Score 47; DB 1; Length 372;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1667 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713

```

```

Db      320 TTTAAAAA..... 366
      |||
RESULT  47
Q12023
ID      Q12023 standard; cDNA; 7722 BP.
AC      Q12023;
DT      13-AUG-1991 (first entry)
DE      Norwalk virus genome sense strand
KW      gastroenteritis; shellfish; Norwalk virus; detection; ss.
OS      Norwalk virus.
FH      Key
FT      cds
      Location/Qualifiers
      4543..4924
      /tag= a
      /product= RNA-dependent RNA polymerase
PN      W09107502-A.
PD      30-MAY-1991.
PF      30-OCT-1990; U06285.
PR      08-NOV-1989; US-433492.
PR      27-APR-1990; US-515993.
PR      27-AUG-1990; US-573509.
PA      (PAU) BAYLOR COLLEGE MED.
PI      Estes MK, Jiang X, Graham DY.
PI      WPI: 91-17811/24.
DR      P-PSDB: R12298.
PT      Double stranded cDNA for Norwalk-type virus detection - comprises
PT      fragments large enough to bind Norwalk or related virus genome
PT      for identification and characterisation
PS      Claim 1: Page 40: 70pp; English.
CC      This sequence may contain an additional guanine between G and C at
CC      nucleotides 5354 and 5355. The source of cDNA was stool samples from
CC      human volunteers administered Norwalk virus. A cDNA library was
CC      prepared from nucleic acids extracted from purified viruses.
CC      Positive Norwalk virus clone pUCNV-953 was identified from an
CC      amplified library (see Q12025): It reacted with post- and pre-
CC      infection stool samples from two volunteers. pUCNV-953 was used
CC      to identify other clones containing fragments of the Norwalk virus
CC      genome. The sequence given here was established in this way.
SO      Sequence 7722 BP: 2187 A; 1790 C; 1879 G; 1866 T;

Query Match      2.7%; Score 47; DB 1; Length 7722;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1667 TTTAAAAA..... 1713
      |||
Db      7639 TTTAAAAA..... 7685

RESULT  48
Q11760/c
ID      Q11760 standard; DNA; 86 BP.
AC      Q11760;
DT      22-JUL-1991 (first entry)
DE      Self-complementary, T7 promoter hairpin-forming sequence #2.
KW      T7 bacteriophage; DNA-dependent RNA polymerase;
KW      target sequence amplification; ss.
OS      Synthetic.
FH      Key
FT      stem_loop
      Location/Qualifiers
      1..86
      /tag= a
      /note= "number of T's in loop is 50"
PN      EP-427073-A.
PD      15-MAY-1991.
PF      27-OCT-1990; 120650.
PR      09-NOV-1989; US-434372.
PR      23-AUG-1990; US-569991.
PA      (MOLE-) MOLECULAR DIAGNOSTI.
PI      Dattagupta N.
PI      WPI: 91-141658/20.
DR      Nucleic acid probe for amplification and detection of target
PT

```

```

PT      sequence - capable of forming ligatable hairpin structured
PT      promoter and transcribing target sequence, is sensitive and
PT      useful in medical diagnosis
PS      Claim 2: Page 11: 15pp; English.
CC      The sequence is an example of a preferred T7 RNA polymerase
CC      promoter for use in the invention. There can be from 2 to 50 T's in
CC      the loop region. A probe sequence is ligated to the 3' end of the
CC      promoter region. Upon hybridisation of the probe to a target sequence
CC      and ligation of the hybridised target sequence to the 5' end of the
CC      hairpin-forming sequence, the target sequence can be transcribed by
CC      T7 RNA polymerase. This allows the target sequence to be amplified
CC      (and detected) using a single oligonucleotide component (c.f. PCR).
CC      See also Q11759 and Q11761-Q11764.
SO      Sequence 86 BP: 12 A; 6 C; 6 G; 62 T;

Query Match      2.7%; Score 47; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1667 TTTAAAAA..... 1713
      |||
Db      71 TTTAAAAA..... 25

RESULT  49
Q15139
ID      Q15139 standard; DNA; 1793 BP.
AC      Q15139;
DT      09-MAR-1992 (first entry)
DE      Clone pTACC6 encoding the tomato ACC synthase.
KW      1-aminocyclopropane-1-carboxylic acid synthase; fruit ripening; ss.
OS      Lycopersicon esculentum c.v. Rutgers.
FH      Key
FT      cds
      Location/Qualifiers
      91..1548
      /tag= a
PN      US7579896-A.
PD      12-NOV-1991.
PF      10-SEP-1990; 579896.
PR      10-SEP-1990; US-579896.
PA      (USDA ) US SEC OF AGRICULTURE.
PI      Theologis A, Sato T.
PI      WPI: 91-368895/50.
DR      P-PSDB: R15506.
PT      DNA encoding ACC synthase - used for control of plant development
PT      and for prodn. of ACC synthase, ethylene and ethanol
PS      Disclosure: Fig 8: 73pp; English.
CC      A lambda gt10 library was constructed using cDNA prepared from polyA
CC      RNA isolated from powdered, frozen fruit (i.e. tomatoes).
CC      Recombinant phage containing inserts were plated, transferred to
CC      a nitrocellulose filter and hybridised to zucchini pACC1 cDNA as
CC      probe (see Q15131). A full-length cDNA from tomato, pTACC6, was
CC      pTACC1 was recovered. Additional clones including pTACC6, were
CC      isolated using the 3'-end of pTACC1 as a probe.
CC      See Q15131-Q15140.
SO      Sequence 1793 BP: 631 A; 275 C; 354 G; 533 T;

Query Match      2.7%; Score 47; DB 1; Length 1793;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1667 TTTAAAAA..... 1713
      |||
Db      1734 TTTAAAAA..... 1780

RESULT  50
Q15022
ID      Q15022 standard; DNA; 1410 BP.
AC      Q15022;
DT      16-MAR-1992 (first entry)
DE      Hyoscyamine 6 beta-hydroxylase gene.
PT

```

KW Scopolamine; 6-beta-hydroxyhyoscyamine; ss.
 OS Hyoscyamus niger.
 FH Key Location/Qualifiers
 FT cds 34..1068
 FT /tag= a
 FT polyA_site 1344..1410
 FT /tag= b
 PN J03247277-A.
 PD 05-NOV-1991.
 PF 26-FEB-1990; 042639.
 PR 26-FEB-1990; JP-042639.
 PA (MITC) MITSUI PETROCHEM IND KK.
 PA (MAXF) MAX FACTOR CO.
 DR WPT: 91-356326/50.
 DR P-PSDB: R15347.
 PI tripane alkaloid synthetase and gene - used for prepn. of
 PI 6-beta-hydroxylase hyoscyamine and scopolamine using
 PI multifunctional enzyme
 PS Claim 1: Fig 12: 12pp: Japanese.
 CC The sequence is that of the hyoscyamine 6 beta-hydroxylase (HH)
 CC gene which encodes a multifunctional enzyme and can be used to prepare
 CC both 6 beta-hydroxyhyoscyamine and scopolamine. Use of recombinant DNA
 CC techniques allows prodn. of large amts. of the protein.
 SO Sequence 1410 BP; 483 A; 236 C; 254 G; 437 T;

Query Match 2.7%; Score 47; DB 1; Length 1410;
 Best Local Similarity 100.0%; Pred. NO. 1.5e-08;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1567 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 1342 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1388

Search completed: August 8, 1999, 03:21:46
 Job time: 1630 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 1999, 02:17:30 : Search time 284.88 seconds
(without alignments)
12027.137 Million cell updates/sec

Title: US-09-104-063-3
Sequence: 1737
1 GAATTCACGTGCTGCGC.....CCGCACACACTGGAATTC 1737

Scoring table: OLIGO_NUC

Searched: 2546578 segs, 986266752 residues

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *
44: gb_est25: *
45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *

54: em_est23: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	589	33.9	589	40	AA935648	AA935648 op27c09.s
2	561	32.3	599	33	AA411265	AA411265 zv24h06.r
3	539	31.0	539	36	AA610463	AA610463 np97b09.s
4	518	29.8	593	35	AA576017	AA576017 nm57d04.s
5	501	28.8	501	33	AA426644	AA426644 zv47h11.s
6	498	28.7	657	47	AI493618	AI493618 ln39e07.x
7	486	28.0	693	36	AA634211	AA634211 ac72b06.s
8	480	27.6	651	36	AA634201	AA634201 ac72a07.s
9	470	27.1	470	33	AA426566	AA426566 zv47h11.r
10	469	27.0	536	38	AA804282	AA804282 nw30h06.s
11	462	26.6	792	38	AA781110	AA781110 aj23e10.s
12	461	26.5	556	38	AA743645	AA743645 ny24g09.s
13	441	25.4	455	44	AI300807	AI300807 gn47b03.x
14	439	25.3	463	45	AI344724	AI344724 gp04h09.x
15	437	25.2	482	41	AI002547	AI002547 og90e01.s
16	435	25.0	524	44	AI249788	AI249788 gx51c10.x
17	428	24.6	428	30	AA259199	AA259199 nc17b08.l
18	427	24.6	473	44	AI300764	AI300764 qn50c09.x
19	423	24.4	535	41	AI022030	AI022030 ow72a11.x
20	421	24.2	583	29	AA148292	AA148292 z045e12.s
21	419	24.1	419	44	AI284206	AI284206 q126h07.x
22	409	23.5	409	34	AA490436	AA490436 aa51e04.r
23	409	23.5	465	45	AI370816	AI370816 ta58c10.x
24	403	23.2	509	38	AA760743	AA760743 n210d09.s
25	403	23.2	417	45	AI358446	AI358446 gx19c11.x
26	401	23.1	401	49	AI660355	AI660355 w62b07.x
27	393	22.6	393	33	AA386000	AA386000 EST99758
28	386	22.2	521	34	AA479467	AA479467 zv17f10.r
29	385	22.2	504	43	AI159856	AI159856 qb51c09.x
30	385	22.2	475	44	AI193140	AI193140 qe40f12.s
31	375	21.6	531	46	AI434652	AI434652 t134d02.x
32	366	21.1	537	34	AA479357	AA479357 zv17f10.s
33	354	20.4	405	50	AI682902	AI682902 wc69e07.x
34	354	20.4	786	37	AA707668	AA707668 z129f12.s
35	347	19.7	454	31	AA284857	AA284857 z122e05.s
36	342	19.6	484	31	AA284857	AA284857 z122e05.s
37	328	18.9	484	31	AA284857	AA284857 z122e05.s
38	308	17.7	308	34	AA501992	AA501992 ne51h08.s
39	299	17.2	361	48	AI610565	AI610565 tp41d08.x
40	298	17.2	407	21	T62636	T62636 yc03f11.s1
41	296	17.0	342	36	AA621854	AA621854 ng19e01.s
42	295	17.0	488	33	AA386001	AA386001 EST98759
43	293	16.9	456	32	AA354393	AA354393 EST62864
44	292	16.8	555	38	AA747545	AA747545 nx85d08.s
45	291	16.8	524	48	AI583169	AI583169 tr98d12.x
46	279	16.1	327	22	R78620	R78620 y174a05.s1
47	277	15.9	277	48	AI582682	AI582682 tn16d04.x
48	275	15.8	427	40	AA934643	AA934643 co07i08.s
49	274	15.8	347	22	R48597	R48597 vj65e07.s1
50	274	15.8	274	48	AI581657	AI581657 as03h04.x
51	271	15.6	273	32	AA339027	AA339027 EST44092
52	270	15.5	318	32	AA378777	AA378777 EST91512
53	264	15.2	373	36	AA661523	AA661523 ns27f07.s
54	263	15.1	388	50	AI695339	AI695339 wa23d09.x
55	259	14.9	259	37	AA732228	AA732228 n295d12.s
56	255	14.7	411	32	AA354912	AA354912 EST63512
57	248	14.3	337	32	AA357257	AA357257 EST65953
58	238	13.7	344	32	AA354044	AA354044 EST62310

59	228	13.1	382	22	R78657	R78657 y174a05.r1	132	49	2.8	190	40	C93155	C93155 C93155 D1ct
60	228	13.1	552	37	AA731086	AA731086 n68a05.s	133	49	2.8	239	42	AA116374	AA116374 uF02c08.y
61	226	13.0	380	32	AA354040	AA354040 EST62304	134	49	2.8	494	43	AA121083	AA121083 u69e12.x
62	224	12.9	533	21	T62491	T62491 yC03f11.r1	135	49	2.8	525	42	AA1173457	AA1173457 u72c10.x
63	219	12.6	331	32	AA353948	AA353948 EST62146	136	49	2.8	131	43	AA1173555	AA1173555 u68b01.x
64	218	12.6	270	32	AA354267	AA354267 EST62592	137	49	2.8	312	43	AA1182513	AA1182513 ue21909.x
65	218	12.6	218	20	D20318	D20318 HUMS01292	138	49	2.8	497	43	AA1029685	AA1029685 A021908.x
66	201	11.6	370	30	AA225739	AA225739 nc17b08.s	139	49	2.8	275	44	AA1269625	AA1269625 qw89a06.x
67	201	11.6	307	20	T29035	T29035 EST65954.Hu	140	49	2.8	443	46	AA1397869	AA1397869 NCSMB2T7
68	199	11.5	435	25	RA68699	RA68699 yj65e02.r1	141	49	2.8	196	48	AA1568114	AA1568114 tp25b05.x
69	192	11.1	398	30	AA248174	AA248174 csg1092.s	142	49	2.8	311	51	AA070779	AA070779 A070779
70	175	10.1	350	23	H29001	H29001 ym31f07.s1	143	48	2.8	433	22	RA42275	RA42275 yf199d08.s1
71	174	10.0	492	29	AA129609	AA129609 z111910.r1	144	48	2.8	141	24	D19176	D19176 MUGS00442
72	169	9.7	335	35	AA573161	AA573161 mm51f02.s	145	48	2.8	336	25	N67359	N67359 y250e06.s1
73	166	9.6	200	33	AA436258	AA436258 v224b06.s	146	48	2.8	397	26	W36406	W36406 mb83f10.r1
74	152	8.8	357	47	AA168830	AA168830 t142b08.x	147	48	2.8	327	27	AA008853	AA008853 mg96611.r
75	150	8.6	151	41	AI050023	AI050023 an22d06.x	148	48	2.8	379	27	AA023416	AA023416 mb72d07.r
76	128	7.4	437	50	AI675038	AI675038 w222f10.x	149	48	2.8	311	27	W84180	W84180 mT35f10.r1
77	122	7.0	157	32	AA357215	AA357215 EST65925	150	48	2.8	357	28	AA098354	AA098354 mk22e12.r
78	114	6.6	145	22	R23114	R23114 yh27b12.r1	151	48	2.8	335	28	AA099424	AA099424 z18a05.s
79	112	6.4	214	32	AA353051	AA353051 EST61123	152	48	2.8	464	28	AA099909	AA099909 z191d05.s
80	106	6.1	261	32	AA357216	AA357216 EST65926	153	48	2.8	218	28	AA124210	AA124210 mn46a04.r
81	100	5.8	268	32	AA357216	AA357216 EST65926	154	48	2.8	221	28	AA124216	AA124216 md46e03.r
82	100	5.8	100	36	AA634862	AA634862 db29b06.r	155	48	2.8	175	29	AA146535	AA146535 mt07c01.r
83	99	5.7	249	23	H29103	H29103 ym31f07.r1	156	48	2.8	192	29	AA148291	AA148291 z045e12.r
84	99	5.7	223	31	AA299445	AA299445 EST11954	157	48	2.8	469	29	AA178252	AA178252 mt13g07.r
85	96	5.5	411	50	AI683322	AI683322 tw50f10.x	158	48	2.8	564	29	AA192336	AA192336 zp97g07.s
86	96	4.7	214	81	R23115	R23115 yh27b12.s1	159	48	2.8	656	30	AA195940	AA195940 zp99a03.s
87	75	4.3	254	28	AA091989	AA091989 110224.se	160	48	2.8	375	30	AA216813	AA216813 mu86n01.r
88	75	4.3	409	24	N42987	N42987 yy08b05.r1	161	48	2.8	354	30	AA219283	AA219283 qv17a11.s
89	53	3.1	385	30	AA221639	AA221639 mw50f03.r	162	48	2.8	165	30	AA219867	AA219867 mw62d05.r
90	53	3.1	136	39	C84892	C84892 C84892 D1ct	163	48	2.8	309	31	AA220766	AA220766 mw66g11.r
91	53	3.1	331	40	C90379	C90379 C90379 D1ct	164	48	2.8	240	30	AA222610	AA222610 mw70f11.r
92	53	3.1	250	43	AU053302	AU053302 AU053302	165	48	2.8	335	30	AA238825	AA238825 mx92d02.r
93	53	3.1	48	53	HS000060	HS000060 Homo. sap1	166	48	2.8	308	30	AA240821	AA240821 mw24b07.r
94	52	3.0	454	29	AA133892	AA133892 zn85b07.s	167	48	2.8	352	30	AA253985	AA253985 va10b01.r
95	52	3.0	337	30	AA239795	AA239795 mx80a12.r	168	48	2.8	278	30	AA259893	AA259893 va53a05.r
96	51	2.9	344	30	AA259726	AA259726 va47b11.r	169	48	2.8	546	30	AA266979	AA266979 m299b04.r
97	51	2.9	400	34	AA470523	AA470523 nc71g11.s	170	48	2.8	288	30	AA267611	AA267611 mz94d09.r
98	51	2.9	521	36	AA648402	AA648402 ns22d04.s	171	48	2.8	295	30	AA268010	AA268010 va89q02.r
99	51	2.9	473	37	AA669094	AA669094 aa81e11.s	172	48	2.8	357	30	AA268643	AA268643 va43f04.r
100	51	2.9	539	38	AA743207	AA743207 ny21a08.s	173	48	2.8	240	30	AA268894	AA268894 va44c05.r
101	51	2.9	261	38	AA766932	AA766932 OC88902.s	174	48	2.8	315	30	AA270150	AA270150 va64h01.r
102	51	2.9	248	38	AA790049	AA790049 vt78g05.r	175	48	2.8	376	30	AA270475	AA270475 va68h02.r
103	51	2.9	235	47	AI473808	AI473808 tm03e12.x	176	48	2.8	213	30	AA272713	AA272713 va48h06.r
104	51	2.9	344	47	AI529465	AI529465 va47b11.y	177	48	2.8	241	30	AA272932	AA272932 va39e07.r
105	51	2.9	419	49	AI624950	AI624950 ts48c11.x	178	48	2.8	175	31	AA276464	AA276464 vc45e01.r
106	51	2.9	132	49	AI642594	AI642594 vv57e01.x	179	48	2.8	447	31	AA285502	AA285502 vb87d09.r
107	50	2.9	563	29	AA131231	AA131231 z040h01.r	180	48	2.8	421	31	AA285554	AA285554 vb86c12.r
108	50	2.9	66	30	AA215584	AA215584 zt96c06.s	181	48	2.8	338	31	AA286632	AA286632 vb87a08.r
109	50	2.9	265	30	AA244683	AA244683 mx03f02.r	182	48	2.8	221	31	AA290362	AA290362 vb51g08.r
110	50	2.9	270	30	AA250484	AA250484 mx05f08.r	183	48	2.8	295	33	AA296382	AA296382 vb45h04.r
111	50	2.9	227	34	AA488429	AA488429 ab39e09.s	184	48	2.8	309	33	AA433117	AA433117 vc94h06.r
112	50	2.9	843	35	AA585438	AA585438 TH842.HTC	185	48	2.8	503	33	AA434832	AA434832 vc17a12.r
113	50	2.9	507	38	AA743134	AA743134 vt20g10.s	186	48	2.8	578	34	AA452150	AA452150 vc115a09.r
114	50	2.9	93	42	AI096771	AI096771 q806h07.x	187	48	2.8	285	34	AA4522107	AA4522107 vc116h06.r
115	50	2.9	475	42	AI154860	AI154860 ud80d11.r	188	48	2.8	100	34	AA4529314	AA4529314 vc139b04.r
116	50	2.9	719	49	AF062726	AF062726 AF062726	189	48	2.8	305	35	AA458149	AA458149 vc157a06.r
117	50	2.9	627	53	HS000745	HS000745 Homo. sap1	190	48	2.8	395	35	AA455111	AA455111 vc140g03.s
118	49	2.8	134	28	AA118047	AA118047 mm12d11.r	191	48	2.8	126	35	AA457176	AA457176 vc006e08.r
119	49	2.8	358	30	AA222538	AA222538 mw20f07.r	192	48	2.8	343	35	AA4592646	AA4592646 vc036g05.r
120	49	2.8	229	30	AA254025	AA254025 mv10g02.r	193	48	2.8	309	35	AA4592692	AA4592692 vc029a05.r
121	49	2.8	384	30	AA269935	AA269935 va64d09.r	194	48	2.8	252	35	AA4592892	AA4592892 vc037h11.r
122	49	2.8	703	35	AA588475	AA588475 vc153e08.r	195	48	2.8	282	35	AA4593292	AA4593292 vc057a12.s
123	49	2.8	780	35	AA585453	AA585453 TH508.HTC	196	48	2.8	416	35	AA4597160	AA4597160 vc039g10.r
124	49	2.8	302	36	AA607600	AA607600 v048e06.r	197	48	2.8	240	36	AA607417	AA607417 vc041a11.r
125	49	2.8	474	36	AA611263	AA611263 vc051b10.r	198	48	2.8	355	36	AA611773	AA611773 vc058a04.r
126	49	2.8	280	37	AA676008	AA676008 T2NF0522	199	48	2.8	287	36	AA611938	AA611938 vc086b06.r
127	49	2.8	202	37	AA721469	AA721469 vt27d07.s	200	48	2.8	181	36	AA616528	AA616528 vc050e10.r
128	49	2.8	309	37	AA734289	AA734289 vt26e04.r	201	48	2.8	428	36	AA616585	AA616585 vc065h10.r
129	49	2.8	854	38	AA745985	AA745985 OD18h03.s	202	48	2.8	221	36	AA617580	AA617580 vc177h09.r
130	49	2.8	246	38	AA789783	AA789783 vt79c09.r	203	48	2.8	236	36	AA652574	AA652574 ns73c07.s
131	49	2.8	220	39	AA822122	AA822122 vc35a12.r	204	48	2.8	356	36	AA658033	AA658033 nu15e04.s

C 205	48	400	37	AA65612	AA65612	ag55a06.s	C 278	48	173	46	A1432532	A132532	th39c11.x
C 206	48	362	37	AA673896	AA673896	vo86g04.i	C 279	48	372	46	A1432975	A1432975	th42a06.x
C 207	48	405	37	AA680531	AA680531	t3678 Bio	C 280	48	447	46	A1433657	A1433657	th88a02.x
C 208	48	497	37	AA682049	AA682049	vu13b08.s	C 281	48	480	46	A1443503	A1443503	sa33f04.x
C 209	48	139	37	AA689755	AA689755	vs03h03.i	C 282	48	167	46	A1445620	A1445620	tl08b05.x
C 210	48	462	37	AA715811	AA715811	rw25b10.s	C 283	48	351	46	A1446457	A1446457	tl19e05.x
C 211	48	225	37	AA734226	AA734226	vt26a09.i	C 284	48	200	47	A1467537	A1467537	ve35f05.x
C 212	48	533	37	N96088	N96088	12401 CD4.i	C 285	48	157	47	A1472566	A1472566	tl178b12.x
C 213	48	213	38	AA202239	AA202239	LD02423.5	C 286	48	397	47	A1473347	A1473347	fb58g07.x
C 214	48	357	38	AA744531	AA744531	ny79a06.s	C 287	48	285	47	A1477670	A1477670	tl16g07.x
C 215	48	531	38	AA744713	AA744713	ny67b04.s	C 288	48	293	47	A1480118	A1480118	tl33g03.x
C 216	48	197	38	AA745607	AA745607	nx18c10.s	C 289	48	349	47	A1493858	A1493858	gm51g05.x
C 217	48	282	38	AA748519	AA748519	ny05f06.s	C 290	48	368	47	A1498582	A1498582	tn02e02.x
C 218	48	413	38	AA760851	AA760851	n214f09.s	C 291	48	413	47	A1500658	A1500658	tn93g07.x
C 219	48	354	38	AA761919	AA761919	n242h05.s	C 292	48	275	47	A1503148	A1503148	vm99e10.x
C 220	48	518	38	AA766249	AA766249	oa13c08.s	C 293	48	396	47	A1506810	A1506810	vm61e04.x
C 221	48	389	38	AA766259	AA766259	oa29a06.s	C 294	48	578	47	A1535918	A1535918	v1cpr02.A
C 222	48	258	38	AA778497	AA778497	af20h03.s	C 295	48	526	47	A1537187	A1537187	v1cpr02.A
C 223	48	202	38	AA792161	AA792161	vn71b11.i	C 296	48	118	47	A1537516	A1537516	tp01d01.x
C 224	48	513	38	AA807510	AA807510	ob92a05.s	C 297	48	169	47	A1538342	A1538342	tp04e05.x
C 225	48	599	38	AA807691	AA807691	nw31b08.s	C 298	48	193	47	A1538733	A1538733	tp56d11.x
C 226	48	441	39	AA814485	AA814485	of42a11.s	C 299	48	222	47	A1539028	A1539028	tp57g05.x
C 227	48	168	39	AA815183	AA815183	vp44h07.i	C 300	48	291	47	A1539219	A1539219	tp71d05.x
C 228	48	140	39	AA820047	AA820047	ak01f06 M	C 301	48	188	47	A1540458	A1540458	tp82c01.x
C 229	48	334	39	AA828415	AA828415	oc46a01.s	C 302	48	271	48	A1545505	A1545505	tp84h02.x
C 230	48	348	39	AA830282	AA830282	oc49e02.s	C 303	48	386	48	A1554818	A1554818	tp86d03.x
C 231	48	377	39	AA831447	AA831447	oc74f10.s	C 304	48	279	48	A1558273	A1558273	tp73f12.x
C 232	48	354	39	AA833819	AA833819	od62b10.s	C 305	48	376	48	A1558487	A1558487	tp78a09.x
C 233	48	245	39	AA835970	AA835970	oc80g03.s	C 306	48	302	48	A1558487	A1558487	fb79g10.x
C 234	48	341	39	AA837211	AA837211	od24f08.s	C 307	48	486	48	A1560010	A1560010	tp93b10.x
C 235	48	217	39	AA847452	AA847452	of07e01.s	C 308	48	208	48	A1560184	A1560184	tp95h02.x
C 236	48	531	39	AA870364	AA870364	vq44e05.i	C 309	48	222	48	A1560806	A1560806	tp95h02.x
C 237	48	531	36	AA872185	AA872185	obv78a10.s	C 310	48	383	48	A1564259	A1564259	tp97d10.x
C 238	48	136	39	C83908	C83908	cb83908 D1c	C 311	48	693	48	A15656042	A15656042	tn52f07.x
C 239	48	481	39	C84058	C84058	cb84058 D1c	C 312	48	273	48	A1567501	A1567501	lr89a10.x
C 240	48	212	40	AA904121	AA904121	cg420b11.s	C 313	48	320	48	A1567814	A1567814	lq87a01.x
C 241	48	430	40	AA959983	AA959983	vw53d06.s	C 314	48	230	48	A1568338	A1568338	tn68b10.x
C 242	48	389	40	AA979970	AA979970	MEST4-r3.	C 315	48	337	48	A1577018	A1577018	ti-r-ABO-
C 243	48	262	40	C91078	C91078	cg1078 D1c	C 316	48	351	48	A1584066	A1584066	ts13c06.x
C 244	48	171	40	C91526	C91526	cg1526 D1c	C 317	48	141	48	A1584118	A1584118	ts13h09.x
C 245	48	297	41	A1049575	A1049575	an34b09.x	C 318	48	294	48	A1584353	A1584353	fb92e08.x
C 246	48	426	41	A1049923	A1049923	an34e08.x	C 319	48	146	48	A1590535	A1590535	tw11b06.x
C 247	48	231	42	A1139104	A1139104	qc27g03.x	C 320	48	269	48	A1591566	A1591566	vt25h10.x
C 248	48	111	43	A1224373	A1224373	qx05g05.x	C 321	48	250	48	A1601818	A1601818	fc12c08.x
C 249	48	344	43	AU028297	AU028297	AU028297	C 322	48	402	48	A1609593	A1609593	tw28d05.x
C 250	48	459	43	AU028802	AU028802	AU028802	C 323	48	212	48	A1610822	A1610822	lp38h08.x
C 251	48	446	43	AU031165	AU031165	AU031165	C 324	48	109	49	A1612723	A1612723	lp38h08.x
C 252	48	421	43	C96980	C96980	C96980 C96980 Rice	C 325	48	584	49	A1625421	A1625421	ts68e02.x
C 253	48	379	43	C97078	C97078	cg7078 Rice	C 326	48	76	49	A1628214	A1628214	lv22h05.x
C 254	48	243	43	C97221	C97221	cg7221 Rice	C 327	48	584	49	A1635528	A1635528	ts95c04.x
C 255	48	301	43	C97413	C97413	cg7413 Rice	C 328	48	248	49	A1636187	A1636187	ts95c04.x
C 256	48	515	43	C99321	C99321	cg9321 Rice	C 329	48	354	49	A1636309	A1636309	tz06e07.x
C 257	48	448	43	C99324	C99324	cg9324 Rice	C 330	48	435	49	A1641270	A1641270	tz27f07.x
C 258	48	338	44	A12556335	A12556335	u184h07.x	C 331	48	297	49	A1645693	A1645693	fc21g09.x
C 259	48	523	44	A1262104	A1262104	qz28d10.y	C 332	48	408	49	A1649693	A1649693	ly04c01.x
C 260	48	350	44	A1280517	A1280517	qu03e04.x	C 333	48	308	49	A1653979	A1653979	ly04c01.x
C 261	48	366	44	A1289134	A1289134	gn25b06.x	C 334	48	186	49	A1654015	A1654015	tl04f12.x
C 262	48	215	44	A1289400	A1289400	gw32e06.x	C 335	48	327	49	A1655735	A1655735	tl14h03.x
C 263	48	374	44	A1316658	A1316658	u116g11.y	C 336	48	506	49	A1658794	A1658794	tl03a04.x
C 264	48	115	44	AU037787	AU037787	AU037787	C 337	48	189	49	A1660432	A1660432	we67a11.x
C 265	48	450	45	A1334621	A1334621	lb20e02.x	C 338	48	290	49	AU057778	AU057778	AU057778
C 266	48	206	45	A1335214	A1335214	ta93b05.x	C 339	48	266	49	AU057977	AU057977	AU057977
C 267	48	317	45	A1344124	A1344124	qt26b12.x	C 340	48	323	50	A1656935	A1656935	fc43a09.x
C 268	48	450	45	A1349964	A1349964	ta97h04.x	C 341	48	403	50	A1659805	A1659805	ty30f03.x
C 269	48	390	45	A1351737	A1351737	qt03e04.x	C 342	48	376	50	A1669864	A1669864	wb89g11.x
C 270	48	361	45	A1360377	A1360377	qy85f01.x	C 343	48	370	50	A1671931	A1671931	wc26g07.x
C 271	48	333	45	A1362495	A1362495	qv16h10.x	C 344	48	224	50	A1672764	A1672764	we58e09.x
C 272	48	483	45	A1363744	A1363744	qv78e01.x	C 345	48	340	50	A1674838	A1674838	wc77d07.x
C 273	48	182	45	A1368579	A1368579	qv61b10.x	C 346	48	712	50	A1676881	A1676881	tu58f04.x
C 274	48	277	45	A1371984	A1371984	ta35b10.x	C 347	48	391	50	A1679209	A1679209	tu72f10.x
C 275	48	379	45	A1372009	A1372009	ta35e06.x	C 348	48	220	50	A1679438	A1679438	tu63g11.x
C 276	48	533	46	A1418970	A1418970	tf38c08.x	C 349	48	315	50	A1679529	A1679529	tu64b10.x
C 277	48	351	46	A1427074	A1427074	mn30b04.x	C 350	48	216	50	A1680369	A1680369	tw62f05.x

C 351	48	2.8	405	50	A1683250	C 424	47	2.7	294	28	AA059479	AA059479	zfe6b04.s
C 352	48	2.8	283	50	A1683492	C 425	47	2.7	170	28	AA070777	AA070777	zm54a08.s
C 353	48	2.8	73	50	A1689096	C 426	47	2.7	140	28	AA087985	AA087985	mo39e12.r
C 354	48	2.8	235	50	A1690663	C 427	47	2.7	302	28	AA102339	AA102339	z191c01.s
C 355	48	2.8	249	50	A1692568	C 428	47	2.7	258	28	AA104461	AA104461	mo55b01.r
C 356	48	2.8	356	50	A1692941	C 429	47	2.7	162	28	AA114771	AA114771	ml14e09.r
C 357	48	2.8	366	50	A1694156	C 430	47	2.7	119	28	AA116956	AA116956	me29b03.r
C 358	48	2.8	144	50	A1696583	C 431	47	2.7	244	28	AA117061	AA117061	ml29d08.r
C 359	48	2.8	236	50	A1699255	C 432	47	2.7	142	28	AA117240	AA117240	ml19f12.r
C 360	48	2.8	384	50	A1701890	C 433	47	2.7	164	28	AA117464	AA117464	ml22d09.r
C 361	48	2.8	635	50	A1702071	C 434	47	2.7	432	28	AA119775	AA119775	mp91d12.r
C 362	48	2.8	119	50	A17060519	C 435	47	2.7	134	28	AA120558	AA120558	ml13a03.r
C 363	48	2.8	293	50	F34241	C 436	47	2.7	338	28	AA120562	AA120562	ml13c03.r
C 364	48	2.8	71	50	F37855	C 437	47	2.7	153	28	AA122192	AA122192	zn28e03.s
C 365	48	2.8	395	51	A1711980	C 438	47	2.7	102	28	AA123210	AA123210	ml25902.r
C 366	48	2.8	362	51	A1721560	C 439	47	2.7	202	28	C22148	C22148	ml25902.r
C 367	48	2.8	654	51	A1728331	C 440	47	2.7	259	29	AA137934	AA137934	ml22e10.r
C 368	48	2.8	626	51	A1730900	C 441	47	2.7	388	29	AA143304	AA143304	zo37a02.s
C 369	48	2.8	564	51	A1734555	C 442	47	2.7	226	29	AA143316	AA143316	zo37c02.s
C 370	48	2.8	952	53	HSM000564	C 443	47	2.7	314	29	AA144045	AA144045	me65a08.r
C 371	48	2.8	1023	53	HSM001913	C 444	47	2.7	296	29	AA172704	AA172704	me58b10.r
C 372	48	2.8	1135	53	HSM001933	C 445	47	2.7	368	29	AA172933	AA172933	me77c02.r
C 373	48	2.8	273	53	HSM002678	C 446	47	2.7	384	29	AA177238	AA177238	me78c04.r
C 374	48	2.8	172	53	HSM003081	C 447	47	2.7	491	29	AA179186	AA179186	me84e08.r
C 375	48	2.8	147	53	HSM003359	C 448	47	2.7	793	29	AA182270	AA182270	me88c02.r
C 376	48	2.8	465	53	HSM003763	C 449	47	2.7	286	29	AA182287	AA182287	me83f01.r
C 377	48	2.8	190	53	HSM005078	C 450	47	2.7	213	29	AA183263	AA183263	me76h12.r
C 378	48	2.8	790	53	HSM007703	C 451	47	2.7	247	29	AA183305	AA183305	me77c02.r
C 379	48	2.8	650	54	HSM010334	C 452	47	2.7	313	29	AA183047	AA183047	me78c04.r
C 380	48	2.8	471	54	HSM010561	C 453	47	2.7	219	29	AA184945	AA184945	me84e08.r
C 381	48	2.8	86	54	HSM010597	C 454	47	2.7	302	29	AA184997	AA184997	me88c02.r
C 382	48	2.8	249	54	HSM011323	C 455	47	2.7	354	29	AA185015	AA185015	me88g11.r
C 383	48	2.8	264	54	HSM011474	C 456	47	2.7	239	29	AA185060	AA185060	me89g11.r
C 384	48	2.8	454	54	HSM011933	C 457	47	2.7	304	29	AA185352	AA185352	mu51d02.r
C 385	48	2.8	285	54	HSM012002	C 458	47	2.7	489	29	AA185434	AA185434	me51d02.r
C 386	47	2.7	151	20	T04695	C 459	47	2.7	370	29	AA186025	AA186025	me35e04.r
C 387	47	2.7	135	20	T25552	C 460	47	2.7	381	29	AA186144	AA186144	me71907.r
C 388	47	2.7	175	20	T28421	C 461	47	2.7	171	29	AA186960	AA186960	me71907.r
C 389	47	2.7	413	20	T49724	C 462	47	2.7	155	29	AA193813	AA193813	rs08g10.r
C 390	47	2.7	75	21	R10067	C 463	47	2.7	577	29	AA194303	AA194303	zg04e11.s
C 391	47	2.7	109	21	T69241	C 464	47	2.7	397	29	U78999	U78999	me78999.me
C 392	47	2.7	247	22	H12358	C 465	47	2.7	746	30	AA195998	AA195998	zp99f05.s
C 393	47	2.7	315	22	R36363	C 466	47	2.7	500	30	AA196541	AA196541	zg08e10.s
C 394	47	2.7	228	22	R43224	C 467	47	2.7	342	30	AA198180	AA198180	mv49g09.r
C 395	47	2.7	365	22	R80916	C 468	47	2.7	188	30	AA198269	AA198269	mv47f10.r
C 396	47	2.7	359	23	H42657	C 469	47	2.7	303	30	AA198866	AA198866	mv42c07.r
C 397	47	2.7	159	24	H79905	C 470	47	2.7	419	30	AA199403	AA199403	mv39h11.r
C 398	47	2.7	124	24	N25033	C 471	47	2.7	282	30	AA210320	AA210320	mv72a06.r
C 399	47	2.7	61	24	N33110	C 472	47	2.7	286	30	AA210335	AA210335	mv72h03.r
C 400	47	2.7	59	24	N39043	C 473	47	2.7	194	30	AA210586	AA210586	mu71b03.r
C 401	47	2.7	299	24	N41289	C 474	47	2.7	119	30	AA212045	AA212045	mo90g08.r
C 402	47	2.7	350	25	N43396	C 475	47	2.7	100	30	AA212139	AA212139	mo90d11.r
C 403	47	2.7	294	25	N63391	C 476	47	2.7	505	30	AA214091	AA214091	zn58h02.s
C 404	47	2.7	200	25	N97837	C 477	47	2.7	230	30	AA215267	AA215267	mu97c05.r
C 405	47	2.7	128	25	N98038	C 478	47	2.7	258	30	AA216888	AA216888	mu93c06.r
C 406	47	2.7	276	25	N98166	C 479	47	2.7	351	30	AA216946	AA216946	mv75g10.r
C 407	47	2.7	352	25	N14029	C 480	47	2.7	333	30	AA217935	AA217935	mv56e10.r
C 408	47	2.7	306	26	W18492	C 481	47	2.7	381	30	AA220788	AA220788	mv69f04.r
C 409	47	2.7	355	26	W33314	C 482	47	2.7	379	30	AA222625	AA222625	mv77g03.r
C 410	47	2.7	194	26	W38553	C 483	47	2.7	363	30	AA230383	AA230383	mv73a12.r
C 411	47	2.7	450	26	W52670	C 484	47	2.7	329	30	AA230802	AA230802	mu13e02.r
C 412	47	2.7	285	26	W69041	C 485	47	2.7	327	30	AA237194	AA237194	mv96f08.r
C 413	47	2.7	468	26	W73128	C 486	47	2.7	360	30	AA237826	AA237826	mx29d06.r
C 414	47	2.7	92	26	W74529	C 487	47	2.7	351	30	AA238147	AA238147	mx28b02.r
C 415	47	2.7	329	26	W84077	C 488	47	2.7	240	30	AA241764	AA241764	mw25g09.r
C 416	47	2.7	345	26	W91679	C 489	47	2.7	126	30	AA241897	AA241897	mx22e02.r
C 417	47	2.7	122	27	AA033323	C 490	47	2.7	452	30	AA242021	AA242021	mx22d09.r
C 418	47	2.7	360	27	AA050541	C 491	47	2.7	326	30	AA242422	AA242422	mw26f02.r
C 419	47	2.7	535	27	AA056832	C 492	47	2.7	247	30	AA242422	AA242422	mw26a11.r
C 420	47	2.7	240	27	AA056832	C 493	47	2.7	278	30	AA250004	AA250004	mw59a10.r
C 421	47	2.7	368	27	W82092	C 494	47	2.7	240	30	AA250343	AA250343	mx59b11.r
C 422	47	2.7	382	27	W99042	C 495	47	2.7	304	30	AA253695	AA253695	mx85c03.r
C 423	47	2.7	510	28	AA058861	C 496	47	2.7	301	30	AA254104	AA254104	va08f05.r

497	47	2.7	182	30	AA254324	AA254324	val5d04.r	C 570	47	2.7	261	34	AA481167	AA481167	aa34h09.s
498	47	2.7	217	30	AA254413	AA254413	val5b03.r	C 571	47	2.7	550	34	AA496558	AA496558	vf99g01.s
499	47	2.7	293	30	AA254562	AA254562	val3c06.r	C 572	47	2.7	106	34	AA502794	AA502794	ne43g08.s
500	47	2.7	297	30	AA259398	AA259398	vas1b01.r	C 573	47	2.7	75	34	AA503384	AA503384	ne45b11.s
501	47	2.7	178	30	AA259635	AA259635	va47f03.r	C 574	47	2.7	291	34	AA505511	AA505511	nh84e11.s
502	47	2.7	347	30	AA259742	AA259742	va47f12.r	C 575	47	2.7	315	34	AA505529	AA505529	nh84h02.s
503	47	2.7	127	30	AA260102	AA260102	va52a09.r	C 576	47	2.7	222	34	AA512115	AA512115	vi16b10.r
504	47	2.7	195	30	AA260691	AA260691	va01g04.r	C 577	47	2.7	324	34	AA512963	AA512963	nj16m01.s
505	47	2.7	245	30	AA260795	AA260795	va21g03.r	C 578	47	2.7	355	34	AA512985	AA512985	nj16m01.s
506	47	2.7	227	30	AA261113	AA261113	m265h09.r	C 579	47	2.7	258	34	AA516371	AA516371	nf16g01.s
507	47	2.7	222	30	AA261246	AA261246	va51h04.r	C 580	47	2.7	233	34	AA517983	AA517983	vf11b09.r
508	47	2.7	225	30	AA261593	AA261593	m265a08.r	C 581	47	2.7	88	34	AA518548	AA518548	vi17c09.r
509	47	2.7	338	30	AA265772	AA265772	m273b03.r	C 582	47	2.7	598	34	AA521299	AA521299	aa79f05.s
510	47	2.7	368	30	AA266373	AA266373	m267a05.r	C 583	47	2.7	268	34	AA521767	AA521767	vi12e09.r
511	47	2.7	295	30	AA266600	AA266600	m267f10.r	C 584	47	2.7	410	34	AA522073	AA522073	vi16e05.r
512	47	2.7	239	30	AA266700	AA266700	m298b08.r	C 585	47	2.7	365	34	AA522097	AA522097	vi16g07.r
513	47	2.7	247	30	AA266722	AA266722	m298f05.r	C 586	47	2.7	545	34	AA525540	AA525540	n165b07.s
514	47	2.7	261	30	AA267108	AA267108	m297d09.r	C 587	47	2.7	408	34	AA525735	AA525735	tenf0042
515	47	2.7	400	30	AA267358	AA267358	m291d09.r	C 588	47	2.7	319	34	AA529372	AA529372	v139e03.r
516	47	2.7	295	30	AA267875	AA267875	va18g08.r	C 589	47	2.7	532	34	AA529509	AA529509	vi40f02.r
517	47	2.7	326	30	AA268028	AA268028	vb08h07.r	C 590	47	2.7	284	34	AA529805	AA529805	
518	47	2.7	356	30	AA268084	AA268084	va89h01.r	C 591	47	2.7	232	34	AA529859	AA529859	
519	47	2.7	380	30	AA268844	AA268844	va41d11.r	C 592	47	2.7	523	34	C24342	C24342	
520	47	2.7	301	30	AA268800	AA268800	va43d08.r	C 593	47	2.7	241	34	C24389	C24389	
521	47	2.7	359	30	AA268972	AA268972	va99a06.r	C 594	47	2.7	332	34	C24395	C24395	
522	47	2.7	287	30	AA269333	AA269333	va41b08.r	C 595	47	2.7	70	35	AA537101	AA537101	vj96g06.r
523	47	2.7	213	30	AA269395	AA269395	va56f12.r	C 596	47	2.7	340	35	AA538496	AA538496	vj53h07.r
524	47	2.7	401	30	AA269763	AA269763	va63d03.r	C 597	47	2.7	352	35	AA538516	AA538516	vj01b06.r
525	47	2.7	327	30	AA269806	AA269806	va64h11.r	C 598	47	2.7	653	35	AA538551	AA538551	vj54b09.r
526	47	2.7	196	30	AA269952	AA269952	va59a11.r	C 599	47	2.7	675	35	AA541878	AA541878	vj54h03.r
527	47	2.7	320	30	AA270173	AA270173	va63d08.r	C 600	47	2.7	704	35	AA541933	AA541933	vj58g03.r
528	47	2.7	286	30	AA270284	AA270284	va63d12.r	C 601	47	2.7	716	35	AA541939	AA541939	vj58c11.r
529	47	2.7	385	30	AA270384	AA270384	vb72e11.r	C 602	47	2.7	271	35	AA543164	AA543164	vi18j06.r
530	47	2.7	403	30	AA270895	AA270895	va70d10.r	C 603	47	2.7	710	35	AA545680	AA545680	vk01g01.r
531	47	2.7	319	30	AA270951	AA270951	vb79d11.r	C 604	47	2.7	552	35	AA552445	AA552445	nk15d11.s
532	47	2.7	298	30	AA271200	AA271200	va73c02.r	C 605	47	2.7	249	35	AA555735	AA555735	vk10c10.r
533	47	2.7	367	30	AA271960	AA271960	va79g05.r	C 606	47	2.7	499	35	AA557132	AA557132	n17a12.s
534	47	2.7	267	30	AA271992	AA271992	va80a06.r	C 607	47	2.7	104	35	AA560851	AA560851	vk94f08.r
535	47	2.7	415	30	AA272436	AA272436	va42a07.r	C 608	47	2.7	224	35	AA562295	AA562295	vk97a12.r
536	47	2.7	311	30	AA272760	AA272760	vc01a03.r	C 609	47	2.7	243	35	AA571655	AA571655	vm09e04.r
537	47	2.7	306	30	AA273250	AA273250	vc01d10.r	C 610	47	2.7	111	35	AA574881	AA574881	vm33a12.r
538	47	2.7	364	30	AA273444	AA273444	vb99f04.r	C 611	47	2.7	274	35	AA575541	AA575541	vm03c10.r
539	47	2.7	314	30	AA273940	AA273940	vb96f12.r	C 612	47	2.7	354	35	AA575800	AA575800	vn63a02.r
540	47	2.7	30	AA274084	AA274084	vc22775	C 613	47	2.7	451	35	AA575922	AA575922	nm56a11.s	
541	47	2.7	1104	30	C22775	AA275358	vc07d01.r	C 614	47	2.7	87	35	AA579618	AA579618	nm79c06.s
542	47	2.7	227	31	AA275358	AA275358	vc14f09.r	C 615	47	2.7	328	35	AA583169	AA583169	nm37d01.s
543	47	2.7	427	31	AA275763	AA275763	vc25g06.r	C 616	47	2.7	102	35	AA587120	AA587120	nm70b05.s
544	47	2.7	242	31	AA275879	AA275879	vb91b07.r	C 617	47	2.7	185	35	AA589036	AA589036	vi165d01.r
545	47	2.7	238	31	AA277536	AA277536	zs84a11.s	C 618	47	2.7	222	35	AA589218	AA589218	vk24f01.r
546	47	2.7	368	31	AA279293	AA279293	va67a10.r	C 619	47	2.7	146	35	AA590017	AA590017	vm22e09.r
547	47	2.7	303	31	AA285391	AA285391	vb90h10.r	C 620	47	2.7	270	35	AA590062	AA590062	vn52d05.r
548	47	2.7	271	31	AA285459	AA285459	vc32c04.r	C 621	47	2.7	248	35	AA590219	AA590219	vn56f12.r
549	47	2.7	210	31	AA285469	AA285469	vb82g08.r	C 622	47	2.7	365	35	AA590672	AA590672	vn58b12.r
550	47	2.7	319	31	AA285492	AA285492	vb82g08.r	C 623	47	2.7	357	35	AA590688	AA590688	vn58g11.r
551	47	2.7	263	31	AA285774	AA285774	vb82g08.r	C 624	47	2.7	202	35	AA590931	AA590931	vm21h01.r
552	47	2.7	416	31	AA288487	AA288487	vb39g10.r	C 625	47	2.7	326	35	AA591006	AA591006	vn61f03.r
553	47	2.7	197	31	AA290346	AA290346	vc33c09.r	C 626	47	2.7	197	35	AA592391	AA592391	vo35a09.r
554	47	2.7	155	32	AA355098	AA355098	EST63446.r	C 627	47	2.7	136	35	AA596390	AA596390	vo38c07.r
555	47	2.7	352	33	AA390106	AA390106	vb28d08.r	C 628	47	2.7	326	35	AA597246	AA597246	
556	47	2.7	357	33	AA390238	AA390238	CPEST.126	C 629	47	2.7	351	35	C24633	C24633	vo28d05.r
557	47	2.7	307	33	AA396391	AA396391	vb40h11.r	C 630	47	2.7	282	35	C24639	C24639	vo28d05.r
558	47	2.7	264	33	AA397033	AA397033	mx84f09.r	C 631	47	2.7	255	35	C24652	C24652	
559	47	2.7	243	33	AA414924	AA414924	vc50g04.r	C 632	47	2.7	288	35	C24675	C24675	
560	47	2.7	590	33	AA415124	AA415124	MG0087.RC	C 633	47	2.7	498	35	C25597	C25597	
561	47	2.7	329	33	AA427328	AA427328	ve24b05.r	C 634	47	2.7	504	35	C25617	C25617	
562	47	2.7	203	33	AA449768	AA449768	zx07h01.s	C 635	47	2.7	412	35	C25619	C25619	
563	47	2.7	284	33	AA450824	AA450824	vf78h05.r	C 636	47	2.7	510	35	C25685	C25685	
564	47	2.7	539	33	C23689	C23689	C33688.Dict	C 637	47	2.7	91	35	C25755	C25755	
565	47	2.7	193	33	C23733	C23733	C3733.Dict	C 638	47	2.7	363	35	C298527	C298527	
566	47	2.7	268	34	AA472895	AA472895	vb30g12.r	C 639	47	2.7	332	36	AA602414	AA602414	no30f05.s
567	47	2.7	523	34	AA473081	AA473081	vd44b07.r	C 640	47	2.7	395	36	AA604005	AA604005	no47d12.s
568	47	2.7	558	34	AA479803	AA479803	zu43a07.r	C 641	47	2.7	305	36	AA606362	AA606362	vo47g09.r
569	47	2.7	256	34	AA481053	AA481053	aa29d04.s	C 642	47	2.7	321	36	AA606942	AA606942	vm39d12.r

643	47	2.7	231	36	AA607037	vm95b06.r	C 716	47	2.7	565	38	AA748729	ny02h03.s
644	47	2.7	59	36	AA607050	vm95d06.r	C 717	47	2.7	428	38	AA749349	AA749349 nx99f03.s
645	47	2.7	264	36	AA607075	vm95h03.r	C 718	47	2.7	591	38	AA752896	AA752896 97AS0589
646	47	2.7	142	36	AA607117	vm92b03.r	C 719	47	2.7	292	38	AA753389	AA753389 97BS0012
647	47	2.7	271	36	AA607299	vm96d09.r	C 720	47	2.7	656	38	AA753511	AA753511 96BS0678
648	47	2.7	244	36	AA607418	vm04a12.r	C 721	47	2.7	490	38	AA753566	AA753566 97BS0038
649	47	2.7	131	36	AA607427	vm04i11.r	C 722	47	2.7	532	38	AA753611	AA753611 97BS0091
650	47	2.7	100	36	AA607932	vm39e04.r	C 723	47	2.7	548	38	AA753634	AA753634 97BS0116
651	47	2.7	258	36	AA608073	vm29e10.r	C 724	47	2.7	608	38	AA753645	AA753645 97BS0127
652	47	2.7	167	36	AA608256	vm61b09.r	C 725	47	2.7	561	38	AA753690	AA753690 97BS0181
653	47	2.7	419	36	AA608276	vm61g11.r	C 726	47	2.7	526	38	AA753741	AA753741 97BS0242
654	47	2.7	127	36	AA608316	vm85e03.r	C 727	47	2.7	591	38	AA753847	AA753847 97BS0279
655	47	2.7	331	36	AA611505	vm08b08.r	C 728	47	2.7	592	38	AA753866	AA753866 97BS0307
656	47	2.7	270	36	AA611723	vm08f08.r	C 729	47	2.7	488	38	AA753896	AA753896 97BS0342
657	47	2.7	377	36	AA611806	vm03f06.r	C 730	47	2.7	578	38	AA753952	AA753952 97BS0408
C 658	47	2.7	498	36	AA614244	np09g04.s	C 731	47	2.7	175	38	AA755151	AA755151 v91a10.r
659	47	2.7	345	36	AA616335	vm65a10.r	C 732	47	2.7	323	38	AA755717	AA755717 vm12h11.r
660	47	2.7	397	36	AA616794	vm68b03.r	C 733	47	2.7	197	38	AA756676	AA756676 vm56g09.r
661	47	2.7	202	36	AA617056	vm12c06.r	C 734	47	2.7	339	38	AA760884	AA760884 n218d01.s
662	47	2.7	171	36	AA617507	v177b02.r	C 735	47	2.7	483	38	AA761545	AA761545 n223a09.s
663	47	2.7	340	36	AA619045	vm67a12.r	C 736	47	2.7	412	38	AA761573	AA761573 n23d08.s
C 664	47	2.7	470	36	AA622416	nm045f02.s	C 737	47	2.7	240	38	AA761608	AA761608 n28b05.s
665	47	2.7	397	36	AA624123	vm075h03.s	C 738	47	2.7	325	38	AA764206	AA764206 vm45e10.r
666	47	2.7	262	36	AA638541	vm054e02.r	C 739	47	2.7	341	38	AA764377	AA764377 vp08c12.r
667	47	2.7	245	36	AA638677	vm056b09.r	C 740	47	2.7	365	38	AA764946	AA764946 n260f01.s
668	47	2.7	278	36	AA638755	vm094h11.r	C 741	47	2.7	140	38	AA765198	AA765198 n262b04.s
C 669	47	2.7	136	36	AA641818	nm19f08.s	C 742	47	2.7	344	38	AA765346	AA765346 nm06d06.s
C 670	47	2.7	557	36	AA643953	np042h06.s	C 743	47	2.7	411	38	AA765751	AA765751 nm07g08.s
C 671	47	2.7	499	36	AA648480	nm33d10.s	C 744	47	2.7	405	38	AA765849	AA765849 nm25f06.s
C 672	47	2.7	447	36	AA651857	nm38f01.s	C 745	47	2.7	224	38	AA767009	AA767009 nm42a03.s
C 673	47	2.7	580	36	AA653370	nm655f07.s	C 746	47	2.7	103	38	AA767924	AA767924 nm06a03.s
C 674	47	2.7	509	36	AA653460	nm66c12.s	C 747	47	2.7	473	38	AA782332	AA782332 ac26e08.s
C 675	47	2.7	387	36	AA653587	nm61g09.s	C 748	47	2.7	284	38	AA786399	AA786399 15d01a1.f
C 676	47	2.7	323	36	AA662155	nm66b09.s	C 749	47	2.7	366	38	AA790030	AA790030 vm80e06.r
C 677	47	2.7	360	36	C62130	C62130 yu11	C 750	47	2.7	190	38	AA793183	AA793183 vp29e06.r
678	47	2.7	479	37	AA673883	vm086f02.r	C 751	47	2.7	268	38	AA794469	AA794469 vm66d04.r
679	47	2.7	308	37	AA674843	vm51a05.r	C 752	47	2.7	148	38	AA794691	AA794691 vm63c05.r
680	47	2.7	304	37	AA674852	vm51b02.r	C 753	47	2.7	380	38	AA795084	AA795084 vm09f11.r
681	47	2.7	258	37	AA674905	vm57c10.r	C 754	47	2.7	266	38	AA795884	AA795884 vt28g06.r
682	47	2.7	201	37	AA689105	vm06f09.r	C 755	47	2.7	434	38	AA796766	AA796766 vp27c11.r
C 683	47	2.7	451	37	AA689584	nm66a03.s	C 756	47	2.7	609	38	AA796779	AA796779 vp31a05.r
684	47	2.7	413	37	AA691245	vm14d03.r	C 757	47	2.7	211	38	AA796787	AA796787 vp31b04.r
685	47	2.7	401	37	AA692516	vm19f12.r	C 758	47	2.7	378	38	AA796807	AA796807 vp31d09.r
686	47	2.7	254	37	AA697665	HL02953.5	C 759	47	2.7	222	38	AA797898	AA797898 vm33g04.r
687	47	2.7	557	37	AA701785	PM1FG_139	C 760	47	2.7	252	38	AA804541	AA804541 nm28d07.s
688	47	2.7	623	37	AA701797	PM1FG_150	C 761	47	2.7	284	38	AA804843	AA804843 nm44e06.s
689	47	2.7	330	37	AA709867	vm35a08.r	C 762	47	2.7	379	38	AA805434	AA805434 nm05404.s
690	47	2.7	276	37	AA710288	vm13a01.r	C 763	47	2.7	337	38	AA805964	AA805964 nm05c05.s
691	47	2.7	384	37	AA710451	vm42f07.r	C 764	47	2.7	435	38	AA806378	AA806378 nm05g02.s
692	47	2.7	175	37	AA710467	vm42h10.r	C 765	47	2.7	218	38	AA806605	AA806605 nm06g03.s
693	47	2.7	240	37	AA710521	vm10f03.r	C 766	47	2.7	297	38	AA806720	AA806720 nm07e04.s
694	47	2.7	357	37	AA710569	vm14c05.r	C 767	47	2.7	427	38	AA806757	AA806757 nm07f06.s
695	47	2.7	312	37	AA711183	vm56g09.r	C 768	47	2.7	437	38	AA807363	AA807363 nm08d04.s
C 696	47	2.7	165	37	AA713781	nm70a12.s	C 769	47	2.7	269	38	AA808112	AA808112 nm08f07.s
697	47	2.7	382	37	AA727832	vm33d10.r	C 770	47	2.7	133	38	AA808175	AA808175 nm09h12.s
698	47	2.7	367	37	AA727836	vm33e02.r	C 771	47	2.7	308	38	AA810226	AA810226 nm14a12.s
699	47	2.7	391	37	AA727881	vm34a11.r	C 772	47	2.7	437	38	AA810677	AA810677 nm17a09.s
C 700	47	2.7	391	37	AA731184	nm56e12.s	C 773	47	2.7	474	38	AA811284	AA811284 nm06f04.s
C 701	47	2.7	436	37	AA731241	nm59f04.s	C 774	47	2.7	361	39	AA814343	AA814343 nm06g07.s
C 702	47	2.7	151	37	AA731711	nm70c03.s	C 775	47	2.7	320	39	AA814517	AA814517 nm06d12.s
C 703	47	2.7	618	38	AA738219	nm17c11.s	C 776	47	2.7	185	39	AA814691	AA814691 nm06e12.s
C 704	47	2.7	463	38	AA738499	NCP4G2T7	C 777	47	2.7	206	39	AA815526	AA815526 nm07e02.r
C 705	47	2.7	263	38	AA740450	nm24b08.s	C 778	47	2.7	639	39	AA816049	AA816049 nm07f02.r
C 706	47	2.7	377	38	AA741286	nm20d01.s	C 779	47	2.7	508	39	AA822925	AA822925 nm07h11.r
C 707	47	2.7	203	38	AA745012	nm20c12.s	C 780	47	2.7	563	39	AA822944	AA822944 nm07g09.r
C 708	47	2.7	294	38	AA745048	nm16b05.s	C 781	47	2.7	299	39	AA823385	AA823385 nm07a08.r
C 709	47	2.7	347	38	AA745387	nm23a01.s	C 782	47	2.7	335	39	AA830396	AA830396 nm05g11.s
C 710	47	2.7	933	38	AA744557	nm79d07.s	C 783	47	2.7	286	39	AA830421	AA830421 nm05a01.s
C 711	47	2.7	411	38	AA744668	nm73g12.s	C 784	47	2.7	345	39	AA830431	AA830431 nm05b05.s
C 712	47	2.7	330	38	AA745069	nm73g12.s	C 785	47	2.7	299	39	AA830709	AA830709 nm05h06.s
C 713	47	2.7	607	38	AA745959	nm18e09.s	C 786	47	2.7	473	39	AA830815	AA830815 nm05e10.s
C 714	47	2.7	395	38	AA748697	nm57h06.s	C 787	47	2.7	332	39	AA831948	AA831948 nm05d02.s
C 715	47	2.7	406	38	AA748698	nm57h07.s	C 788	47	2.7	223	39	AA832453	AA832453 nm05g11.s

789	47	2.7	331	39	AA833305	AA833305	od65e08.r	862	47	2.7	450	40	C89917	C89917	Dict
C 750	47	2.7	440	39	AA834534	AA834534	od65a06.s	863	47	2.7	172	40	C89934	C89934	Dict
C 791	47	2.7	242	35	AA835168	AA835168	od16e07.s	864	47	2.7	241	40	C90027	C90027	Dict
C 752	47	2.7	247	39	AA836471	AA836471	od38b12.s	865	47	2.7	270	40	C90071	C90071	Dict
C 753	47	2.7	355	39	AA837391	AA837391	od41e12.s	866	47	2.7	270	40	C90078	C90078	Dict
C 794	47	2.7	517	39	AA837930	AA837930	oe93a09.s	867	47	2.7	217	40	C90121	C90121	Dict
C 795	47	2.7	298	39	AA838319	AA838319	oe93b04.s	868	47	2.7	301	40	C90156	C90156	Dict
C 796	47	2.7	409	39	AA847917	AA847917	od39f01.s	869	47	2.7	621	40	C90406	C90406	Dict
C 797	47	2.7	781	39	AA849835	AA849835	EST192f02	870	47	2.7	196	40	C90466	C90466	Dict
C 798	47	2.7	483	39	AA857969	AA857969	oe33h06.s	871	47	2.7	449	40	C90501	C90501	Dict
C 799	47	2.7	198	39	AA862606	AA862606	Oh07b04.s	872	47	2.7	385	40	C90503	C90503	Dict
C 800	47	2.7	204	39	AA862660	AA862660	Oh45c03.s	873	47	2.7	442	40	C90506	C90506	Dict
C 801	47	2.7	329	39	AA871153	AA871153	vg31g07.r	874	47	2.7	216	40	C90718	C90718	Dict
C 802	47	2.7	509	39	AA871931	AA871931	vg43g08.r	875	47	2.7	604	40	C90785	C90785	Dict
C 803	47	2.7	210	39	AA872507	AA872507	o10e11.s	876	47	2.7	264	40	C90816	C90816	Dict
C 804	47	2.7	220	39	AA881672	AA881672	vx21d09.r	877	47	2.7	374	40	C90828	C90828	Dict
C 805	47	2.7	574	39	AA897798	AA897798	NCP4H517	878	47	2.7	403	40	C90829	C90829	Dict
C 806	47	2.7	631	39	AA897888	AA897888	NCC3C977	879	47	2.7	198	40	C90840	C90840	Dict
C 807	47	2.7	409	39	AA897891	AA897891	NCC4A677	880	47	2.7	686	40	C90855	C90855	Dict
C 808	47	2.7	586	39	AA897904	AA897904	NCP4G117	881	47	2.7	348	40	C90856	C90856	Dict
C 809	47	2.7	510	39	AA897913	AA897913	NCC1F177	882	47	2.7	340	40	C90858	C90858	Dict
C 810	47	2.7	480	39	AA897914	AA897914	NCC1F127	883	47	2.7	373	40	C90878	C90878	Dict
C 811	47	2.7	574	39	AA897926	AA897926	NCC2E877	884	47	2.7	390	40	C90883	C90883	Dict
C 812	47	2.7	547	39	AA898229	AA898229	NCC2B977	885	47	2.7	309	40	C90935	C90935	Dict
C 813	47	2.7	557	39	AA898385	AA898385	NCC6D477	886	47	2.7	308	40	C90985	C90985	Dict
C 814	47	2.7	546	39	AA898463	AA898463	NCM5H127	887	47	2.7	256	40	C90988	C90988	Dict
C 815	47	2.7	493	39	AA898589	AA898589	NCC3D177	888	47	2.7	643	40	C90989	C90989	Dict
C 816	47	2.7	586	39	AA898674	AA898674	NCM7G777	889	47	2.7	521	40	C91004	C91004	Dict
C 817	47	2.7	513	39	AA901562	AA901562	NCP6A117	890	47	2.7	276	40	C91123	C91123	Dict
C 818	47	2.7	491	39	AA901781	AA901781	NCP3F477	891	47	2.7	485	40	C91125	C91125	Dict
C 819	47	2.7	523	39	AA901860	AA901860	NCP2E677	892	47	2.7	477	40	C91130	C91130	Dict
C 820	47	2.7	624	39	AA901889	AA901889	NCP2E677	893	47	2.7	252	40	C91203	C91203	Dict
C 821	47	2.7	647	39	AA901984	AA901984	NCP2E677	894	47	2.7	405	40	C91306	C91306	Dict
C 822	47	2.7	264	39	C83835	C83835	Dict	895	47	2.7	267	40	C91337	C91337	Dict
C 823	47	2.7	423	39	C83840	C83840	Dict	896	47	2.7	166	40	C91511	C91511	Dict
C 824	47	2.7	357	39	C83928	C83928	Dict	897	47	2.7	430	40	C91532	C91532	Dict
C 825	47	2.7	211	39	C84164	C84164	Dict	898	47	2.7	194	40	C92022	C92022	Dict
C 826	47	2.7	166	39	C84166	C84166	Dict	899	47	2.7	298	40	C92188	C92188	Dict
C 827	47	2.7	350	39	C84677	C84677	Dict	900	47	2.7	630	40	C92364	C92364	Dict
C 828	47	2.7	331	39	C84704	C84704	Dict	901	47	2.7	607	40	C92385	C92385	Dict
C 829	47	2.7	272	39	C84782	C84782	Dict	902	47	2.7	277	40	C92705	C92705	Dict
C 830	47	2.7	235	39	C84788	C84788	Dict	903	47	2.7	414	40	C92768	C92768	Dict
C 831	47	2.7	578	39	C84827	C84827	Dict	904	47	2.7	362	40	C92779	C92779	Dict
C 832	47	2.7	341	39	C84867	C84867	Dict	905	47	2.7	225	40	C93081	C93081	Dict
C 833	47	2.7	149	40	AA908294	AA908294	oq33f11.s	906	47	2.7	187	40	C93096	C93096	Dict
C 834	47	2.7	504	40	AA908482	AA908482	oq82c12.s	907	47	2.7	216	40	C93106	C93106	Dict
C 835	47	2.7	81	40	AA910956	AA910956	oX85h11.s	908	47	2.7	216	40	C93106	C93106	Dict
C 836	47	2.7	195	40	AA911767	AA911767	oq19c07.s	909	47	2.7	625	40	C93193	C93193	Dict
C 837	47	2.7	469	40	AA915056	AA915056	vz01h11.r	910	47	2.7	343	40	C93249	C93249	Dict
C 838	47	2.7	172	40	AA916133	AA916133	oq32b09.s	911	47	2.7	347	40	C93297	C93297	Dict
C 839	47	2.7	176	40	AA923096	AA923096	oK91c09.s	912	47	2.7	108	41	A1010522	A1010522	Dict
C 840	47	2.7	95	40	AA928539	AA928539	om73g01.s	913	47	2.7	294	41	A1011635	A1011635	Dict
C 841	47	2.7	426	40	AA933083	AA933083	om85f10.s	914	47	2.7	227	41	A1019644	A1019644	Dict
C 842	47	2.7	205	40	AA937566	AA937566	oq79d12.s	915	47	2.7	393	41	A1025513	A1025513	Dict
C 843	47	2.7	361	40	AA937574	AA937574	oq79f06.s	916	47	2.7	455	41	A1039104	A1039104	Dict
C 844	47	2.7	465	40	AA938181	AA938181	oC10d08.s	917	47	2.7	509	41	A1041011	A1041011	Dict
C 845	47	2.7	491	40	AA939876	AA939876	vz95h02.r	918	47	2.7	336	41	A1049659	A1049659	Dict
C 846	47	2.7	333	40	AA945339	AA945339	EST200838	919	47	2.7	71	41	A1050086	A1050086	Dict
C 847	47	2.7	127	40	AA946418	AA946418	EST201917	920	47	2.7	428	41	A1050084	A1050084	Dict
C 848	47	2.7	321	40	AA954676	AA954676	om95d10.s	921	47	2.7	423	41	A1051709	A1051709	Dict
C 849	47	2.7	345	40	AA958898	AA958898	ua19f03.r	922	47	2.7	445	41	A1056694	A1056694	Dict
C 850	47	2.7	437	40	AA959695	AA959695	vw56e06.s	923	47	2.7	69	41	A1061180	A1061180	Dict
C 851	47	2.7	343	40	AA960578	AA960578	vw64a12.s	924	47	2.7	279	41	A1061405	A1061405	Dict
C 852	47	2.7	391	40	AA969375	AA969375	om58e04.s	925	47	2.7	384	41	A1065658	A1065658	Dict
C 853	47	2.7	395	40	AA979714	AA979714	TC-Est-29	926	47	2.7	382	41	A1067648	A1067648	Dict
C 854	47	2.7	356	40	AA980543	AA980543	ua42e08.r	927	47	2.7	544	41	C93698	C93698	Dict
C 855	47	2.7	733	40	AA981912	AA981912	ua34c10.r	928	47	2.7	590	41	C93706	C93706	Dict
C 856	47	2.7	733	40	AA986840	AA986840	ue15f01.x	929	47	2.7	258	41	C93712	C93712	Dict
C 857	47	2.7	496	40	C89741	C89741	Dict	930	47	2.7	589	41	C93745	C93745	Dict
C 858	47	2.7	164	40	C89844	C89844	Dict	931	47	2.7	124	41	C93753	C93753	Dict
C 859	47	2.7	587	40	C89862	C89862	Dict	932	47	2.7	576	41	C93788	C93788	Dict
C 860	47	2.7	309	40	C89869	C89869	Dict	933	47	2.7	252	41	C93792	C93792	Dict
C 861	47	2.7	449	40	C89889	C89889	Dict	934	47	2.7	809	41	C93796	C93796	Dict

935	47	2.7	657	41	C93817	C93817	C93817	DICT
936	47	2.7	930	41	C93828	C93828	C93828	DICT
937	47	2.7	486	41	C93852	C93852	DICT	
938	47	2.7	466	41	C93867	C93867	DICT	
939	47	2.7	813	41	C93868	C93868	DICT	
940	47	2.7	348	41	C93889	C93889	DICT	
941	47	2.7	734	41	C93927	C93927	DICT	
942	47	2.7	441	41	C93943	C93943	DICT	
943	47	2.7	609	41	C93945	C93945	DICT	
944	47	2.7	532	41	C94021	C94021	DICT	
945	47	2.7	270	41	C94036	C94036	DICT	
946	47	2.7	205	41	C94125	C94125	DICT	
947	47	2.7	231	41	C94154	C94154	DICT	
948	47	2.7	355	41	C94167	C94167	DICT	
949	47	2.7	330	41	C94178	C94178	DICT	
950	47	2.7	331	41	C94266	C94266	DICT	
951	47	2.7	194	41	C94286	C94286	DICT	
952	47	2.7	416	41	C94331	C94331	DICT	
953	47	2.7	559	41	C94338	C94338	DICT	
954	47	2.7	488	41	C94347	C94347	DICT	
955	47	2.7	540	41	C94396	C94396	DICT	
956	47	2.7	282	41	C94400	C94400	DICT	
957	47	2.7	174	41	C94407	C94407	DICT	
958	47	2.7	218	41	C94439	C94439	DICT	
959	47	2.7	182	41	C94451	C94451	DICT	
960	47	2.7	406	41	C94510	C94510	DICT	
961	47	2.7	300	41	C94516	C94516	DICT	
962	47	2.7	302	41	C95040	C95040	Pig	
963	47	2.7	326	42	A1073383	A1073383	o013d03.x	
964	47	2.7	270	42	A1074030	A1074030	oY66h06.x	
965	47	2.7	400	42	A1075885	A1075885	o25d09.x	
966	47	2.7	405	42	A1076157	A1076157	oY92h04.x	
967	47	2.7	636	42	A1079226	A1079226	o252b02.x	
968	47	2.7	355	42	A1080256	A1080256	o247906.x	
969	47	2.7	291	42	A1081740	A1081740	o063904.s	
970	47	2.7	437	42	A1086783	A1086783	o277c02.x	
971	47	2.7	458	42	A1088789	A1088789	q21c07.x	
972	47	2.7	346	42	A1088929	A1088929	q21902.x	
973	47	2.7	279	42	A1089782	A1089782	q22907.x	
974	47	2.7	416	42	A1091468	A1091468	o019b03.x	
975	47	2.7	361	42	A1095003	A1095003	q214d05.x	
976	47	2.7	320	42	A1095192	A1095192	qY83b09.s	
977	47	2.7	712	42	A1096512	A1096512	q24c12.x	
978	47	2.7	443	42	A1096694	A1096694	q21b08.x	
979	47	2.7	342	42	A1097229	A1097229	q288d09.x	
980	47	2.7	276	42	A1097410	A1097410	q24c10.x	
981	47	2.7	91	42	A1105872	A1105872	CB02a07.L	
982	47	2.7	192	42	A1105948	A1105948	CB03a04.L	
983	47	2.7	223	42	A1105984	A1105984	CB03f10.L	
984	47	2.7	654	42	A1107087	A1107087	GHO6380.5	
985	47	2.7	540	42	A1107657	A1107657	GHO5457.5	
986	47	2.7	537	42	A1108100	A1108100	GHO6842.5	
987	47	2.7	341	42	A1109273	A1109273	GHO8412.5	
988	47	2.7	461	42	A1110228	A1110228	AM052.ANY	
989	47	2.7	530	42	A1113360	A1113360	GHO6685.5	
990	47	2.7	205	42	A1115071	A1115071	u146e09.x	
991	47	2.7	760	42	A1117098	A1117098	ue30f06.Y	
992	47	2.7	785	42	A1117621	A1117621	u148b03.x	
993	47	2.7	564	42	A1125109	A1125109	am66f01.s	
994	47	2.7	365	42	A1125884	A1125884	q26f104.x	
995	47	2.7	492	42	A1128239	A1128239	q23f109.x	
996	47	2.7	309	42	A1129698	A1129698	q235910.x	
997	47	2.7	459	42	A1129719	A1129719	q24e009.x	
998	47	2.7	795	42	A1133715	A1133715	ue24e05.x	
999	47	2.7	329	42	A1137715	A1137715	ue33h05.x	
1000	47	2.7	501	42	A1134068	A1134068	GHI1418.5	

ALIGNMENTS

RESULT 1
AA935648/c

LOCUS	AA935648	589 bp	mRNA	EST	07-JUL-1998
DEFINITION	op27c09.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1578064 3' similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1 HOMOLOG (HUMAN);, mRNA sequence.				
ACCESSION	AA935648				
NID	g3092805				
VERSION	AA935648.1 GI:3092805				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 589)				
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
COMMENT	Unpublished (1997) On Jan 17, 1998 this sequence version replaced gi:1900670. Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert_strausberg@nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (lnl@image.lnl.gov) for further information. Insert Length: 1009 Std Error: 0.00 Seq primer: -40ml3 fwd. Ex from Amersham High quality sequence stop: 511. Location/Qualifiers 1..589 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:1578064" /clone_lid="Soares_NFL_T_GBC_S1" /lab_host="DH10B" /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBH19W, testis NHT, and B-cell NCI-CGAP GCBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."				
BASE COUNT	220 a 116 c 89 g 164 t				
ORIGIN					
Query Match	33.9%; Score 589; DB 40; Length 589;				
Best Local Similarity	100.0%; Pred. No. 4,5e-245;				
Matches	589; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
OY	1087	AAGCAGGTGACATTCATCTGTTCCAGTGTGCTTCAAGTTTCACCTCCAGC	1146		
DB	589	AAGCAGGTGACATTCATCTGTTCCAGTGTGCTTCAAGTTTCACCTCCAGC	530		
OY	1147	TAACACAGTGTAAAGAGCTTTTATACGATTAATACTTTTATAGTTACACAT	1206		
DB	529	TAACACAGTGTAAAGAGCTTTTATACGATTAATACTTTTATAGTTACACAT	470		
OY	1207	TTTTCAGATTAAGAGCTGACCAATATTGATGTTTATGCTGTGGATTTTGT	1266		
DB	469	TTTTCAGATTAAGAGCTGACCAATATTGATGTTTATGCTGTGGATTTTGT	410		
OY	1267	CTTGTTCTTCTTACTTTTGTGAAGTTAATGACTTATTAATAATTTTGT	1326		
DB	409	CTTGTTCTTCTTACTTTTGTGAAGTTAATGACTTATTAATAATTTTGT	350		
OY	1327	TCATATTGATGTGTCTAGGAGAGAGCTGTGGCAAGTTCTTACTGCTGATCTCG	1386		
DB	349	TCATATTGATGTGTCTAGGAGAGAGCTGTGGCAAGTTCTTACTGCTGATCTCG	290		

QY 1387 TGTAGAGCTGTAAAGGAGCACTGAAACATTTCCAGACGCTGTAGTGAATCAGCTAAAGC 1446
|||||
Db 289 TGTAGGAGCTGTAAAGGAGCACTGAAACATTTCCAGACGCTGTAGTGAATCAGCTAAAGC 230
QY 1447 TAGAATGATCCGACGCTGTATGATAGATATCTCTCCATTCGCGGAGCAAGCTTTT 1506
|||||
Db 229 TAGAATGATCCGACGCTGTATGATAGATATCTCTCCATTCGCGGAGCAAGCTTTT 170
QY 1507 CCTGTCTTAAGACGTGATTTTGTCTGTAGAGATGCGACTTATAACCAAGCCCAAGTG 1566
|||||
Db 169 CCTGTCTTAAGACGTGATTTTGTCTGTAGAGATGCGACTTATAACCAAGCCCAAGTG 110
QY 1567 GTATAGAAATGCTGTGTTTTCAGTTTTCAGAGTGGTGTGATTTAGACACCTACAGTGT 1626
|||||
Db 109 GTATAGAAATGCTGTGTTTTCAGTTTTCAGAGTGGTGTGATTTAGACACCTACAGTGT 50
QY 1627 CAGTCTGTATTAGTGTGTTTAAAGTACATGTTAACTTAAAGAA 1675
|||||
Db 49 CAGTCTGTATTAGTGTGTTTAAAGTACATGTTAACTTAAAGAA 1
RESULT 2
AA411265 599 bp mRNA EST 17-MAY-1997
LOCUS z24h06.f1 Soares_NbHMPu.S1 Homo sapiens cDNA clone IMAGE:754619 5'
DEFINITION similar to gb:U06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA411265
NID g2058847
VERSION AA411265.1 GI:2068847
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 599)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Giesel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Schellenberg,K., Stepien,M., Tan,F., Theising,B.,
White,Y., Wyllie,T., Waterston,R. and Wilson,R.
Washu-Merck EST Project 1997
TITLE Unpublished (1997)
JOURNAL On Sep 12, 1996 this sequence version replaced gi:1406946.
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq. primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 506.
Location/Qualifiers
1. 599
/organism="Homo sapiens"
/db_xref="GDB:5977529"
/db_xref="tixon:9606"
/clone_image="754619"
/clone_lib="Soares_NbHMPu.S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH108"
/note="Organ: mixed (see below); Vector: pTT3-Pac
(Pharmacia) with a modified polylinker; Site:1; NOC I;
Site:2; Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2Nb1M, pregnant uterus
Nb1Pu, and fetal heart NbH19M) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR amplified cDNAs from pools of

5,000 clones made from the same 3 libraries. The pools
consisted of T.M.A.G.E. clones 250232-265223,
340488-345479, and 484488-489479."

BASE COUNT 168 a 93 c 117 g 221 t

ORIGIN

Query Match 32.3%; Score 561; DB 33; Length 599;
Best Local Similarity 100.0%; Pred. No. 5.9e-233;
Matches 561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1092 AGGTGACATTCATGTTTTCACAGTCTGAGCTTCAAGTTCACCTCCAGCTACCA 1151
|||||
Db 15 AGGTGACATTCATGTTTTCACAGTCTGAGCTTCAAGTTCACCTCCAGCTACCA 74
QY 1152 CAGATGTAAAGACTTTTATACGATTAATACITTTTATAGTACACATTTTC 1211
|||||
Db 75 CAGATGTAAAGACTTTTATACGATTAATACITTTTATAGTACACATTTTC 134
QY 1212 AGATATAAAGACGTGACCAATATGTTACAGTTTATGCTGTGATTTTGTCTGT 1271
|||||
Db 135 AGATATAAAGACGTGACCAATATGTTACAGTTTATGCTGTGATTTTGTCTGT 194
QY 1272 GTTCTTGTAGTTTGTGAGCTTATGACTTATATATATTTTGTGTTCA 1331
|||||
Db 195 GTTCTTGTAGTTTGTGAGCTTATGACTTATATATATTTTGTGTTCA 254
QY 1332 TTGATGTGTCTAGCAGGACCTGTGCAAGTTCTTGTGTTGTTGTTGTTA 1391
|||||
Db 255 TTGATGTGTCTAGCAGGACCTGTGCAAGTTCTTGTGTTGTTGTTGTTA 314
QY 1392 GGACGTGTAAGAAAGGACGTGACCTGACAGCTGTTAGTGTGTTGTTGTTGTTA 1451
|||||
Db 315 GGACGTGTAAGAAAGGACGTGACCTGACAGCTGTTAGTGTGTTGTTGTTGTTA 374
QY 1452 ATGATCCCGACGTGTTTATGATAGATATCTCCATTCCTGGAAGTTTCTGCT 1511
|||||
Db 375 ATGATCCCGACGTGTTTATGATAGATATCTCCATTCCTGGAAGTTTCTGCT 434
QY 1512 TCTTAAGACGTGATTTTCTGTAGAGATGCGACTTATACCAAGCCCAAGTGTATA 1571
|||||
Db 435 TCTTAAGACGTGATTTTCTGTAGAGATGCGACTTATACCAAGCCCAAGTGTATA 494
QY 1572 GAAATGCTGTTTTCACGTTTTCAGAGTGGTGTGATTTTCAGACCTCAGTGTACATC 1631
|||||
Db 495 GAAATGCTGTTTTCACGTTTTCAGAGTGGTGTGATTTTCAGACCTCAGTGTACATC 554
QY 1632 TTGTATTAGTGTGTTAATAA 1652
|||||
Db 555 TTGTATTAGTGTGTTAATAA 575

RESULT 3
AA610463/c 539 bp mRNA EST 30-OCT-1997
LOCUS np97b09.s1 NCI-CGAP.Lu1 Homo sapiens cDNA clone IMAGE:1142297 3'
DEFINITION similar to gb:U06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA610463
NID g2458891
VERSION AA610463.1 GI:2458891
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 539)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1395020.

Db 338 TTGATGTGTGTAGCAGACACTGTGGCAAGTCTTGTAGTGTCTGTCTGTGTA 279
QY 1392 GGACGTGTACAAAAAGGAACTGAACATTCAGACCGTGTAGTATCATCGTAAAGCTAGAA 1451
Db 278 GGACGTGTACAAAAAGGAACTGAACATTCAGACCGTGTAGTATCATCGTAAAGCTAGAA 219
QY 1452 ATGATCCCGACGTGTGTATGATAGATTAATCTTCATTCCTCCGTGGAGCTTTTCTGT 1511
Db 218 ATGATCCCGACGTGTGTATGATAGATTAATCTTCATTCCTCCGTGGAGCTTTTCTGT 159
QY 1512 TCTTAAGACGTATTTTGTGTAGAGATGGACCTATTAACCAAGCCCAAGTGTATTA 1571
Db 158 TCTTAAGACGTATTTTGTGTAGAGATGGACCTATTAACCAAGCCCAAGTGTATTA 99
QY 1572 GAAATGCTGCTTTTTCAGTTTTCAGAGATGGTGTATTTTCAGACCTACAGTATACAGTC 1631
Db 98 GAAATGCTGCTTTTTCAGTTTTCAGAGATGGTGTATTTTCAGACCTACAGTATACAGTC 39
QY 1632 TTGTATTAGTGTATTAATAAGTACATGTAACTTA 1669
Db 38 TTGTATTAGTGTATTAATAAGTACATGTAACTTA 1

RESULT 5
AA426644/C 501 bp mRNA EST 16-OCT-1997
LOCUS zva47h1.s1 Soares ovary tumor NBHOT Homo sapiens cDNA clone
DEFINITION IMAGE:756837 3' similar to gb:106797 PROBABLE G PROTEIN-COUPLED
RECEPTOR LCRI HOMOLOG (HUMAN)'' mRNA sequence.
ACCESSION AA426644
NID 92107314
VERSION AA426644.1 GI:2107314
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 501)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisler,G., Jost,S.,
Kucaba,T., Lacey,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Scheilenberg,K., Steptoe,M., Tan,F., Theisling,B.,
White,Y., Wyllie,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
TITLE Unpublished (1997)
JOURNAL On Apr 14, 1993 this sequence version replaced gi:693425.
COMMENT

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41m13 fwd. Et from Amersham
High quality sequence stop: 332.
Location/Qualifiers
1..501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7"
/clone="IMAGE:756837"
/clone_lib="Soares ovary tumor NBHOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/note="organ: ovary; Vector: pRT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
TGTTACCAATCTGAAGTGGAGCGCGCGGTTTCTTTTCTTTT 3'),
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pRT73 vector

FEATURES
source

(Pharmacia), library constructed by Bento Soares and
M.Fatima Bonaldo.
BASE COUNT 185 a 99 c 69 g 148 t
ORIGIN

Query Match 28.8%; Score 501; DB 33; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.8e-207;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 TTTTAACTTACATTTTTCAGATATATAAGACAGACCAATATTTGACATTTTAT 1249
Db 501 TTTTAACTTACATTTTTCAGATATATAAGACAGACCAATATTTGACATTTTAT 442
QY 1250 GCTTGTGCAATTTTGTCTTGTCTTCTTTAGTTTGTGAAGTTTAATGACTTATTA 1309
Db 441 GCTTGTGCAATTTTGTCTTGTCTTCTTTAGTTTGTGAAGTTTAATGACTTATTA 382
QY 1310 TATTAATTTTCTTGTCTTATATATGATGTGTCTAGGAGACCTGTGGCCAGTTCT 1369
Db 381 TATTAATTTTCTTGTCTTATATATGATGTGTCTAGGAGACCTGTGGCCAGTTCT 322
QY 1370 AGTTGCTGATGTCTGTGTAGACTGTAGAAAAGGAACTGAACTTCCAGAGCTGT 1429
Db 321 AGTTGCTGATGTCTGTGTAGACTGTAGAAAAGGAACTGAACTTCCAGAGCTGT 262
QY 1430 AGTGAATCAAGTAACTGAAATGATCCCGACCTTTTATGATAGATATCTCCAT 1489
Db 261 AGTGAATCAAGTAACTGAAATGATCCCGACCTTTTATGATAGATATCTCCAT 202
QY 1490 TCCCGTGAAGCTTTTCCCTGTTCTTAAGACGATTTTGTGTAGAGATGGCACTTAT 1549
Db 201 TCCCGTGAAGCTTTTCCCTGTTCTTAAGACGATTTTGTGTAGAGATGGCACTTAT 142
QY 1550 AACCAAAAGCCCAAGTGTATAGAAAATGCTGTTTTCAGTTTTCAGAGTGGTGTAT 1609
Db 141 AACCAAAAGCCCAAGTGTATAGAAAATGCTGTTTTCAGTTTTCAGAGTGGTGTAT 82
QY 1610 TCAGCACTACAGTGTACAGTCTGTATTAAGTTTATTAAGTAAAGTAACTTA 1669
Db 81 TCAGCACTACAGTGTACAGTCTGTATTAAGTTTATTAAGTAAAGTAACTTA 22
QY 1670 AAAAAAAAAAAAAAAAAAAAAA 1690
Db 21 AAAAAAAAAAAAAAAAAAAAAA 1

RESULT 6
A1493618/C 657 bp mRNA EST 30-MAR-1999
LOCUS th39e07.x1 NCI-CGAP Pauli Homo sapiens cDNA clone IMAGE:2120676 3'
DEFINITION similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN)'' mRNA sequence.
ACCESSION A1493618
NID 94394621
VERSION A1493618.1 GI:4394621
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 657)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On May 18, 1998 this sequence version replaced gi:3137281.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bdrrp/image/image.html

Insert Length: 1615 Std Error: 0.00
Seq primer: -400P from Glbco
High quality sequence stop: 410.
Location/Qualifiers

FEATURES

source

1.657
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2120676"
/clone_1id="NCI CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: Salt; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.72 kb. Life Technologies catalog #: 1548-013"

BASE COUNT 232 a 130 c 109 g 184 t 2 others
ORIGIN

Query Match 28.7%; Score 498; DB 47; Length 657;
Best Local Similarity 99.8%; Pred. No. 1.1e-205;
Matches 548; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1129 TCAAGTTTCACTCCAGCTAACACAGATGTAAAGACTTTTATACGATAAATACT 1188
DB 549 TCAAGTTTCACTCCAGCTAACACAGATGTAAAGACTTTTATACGATAAATACT 490
QY 1189 TTTTATAGTACACATTTTTCAGATATAAAGACGTGCCAATATTGACGTTTAT 1248
DB 489 TTTTATAGTACACATTTTTCAGATATAAAGACGTGCCAATATTGACGTTTAT 430
QY 1249 TCGTGTGTGAGTTTGTCTGTTCTTTTATAGTTTGTGAAGTTTAACTGACTTAT 1308
DB 429 TCGTGTGTGAGTTTGTCTGTTCTTTTATAGTTTGTGAAGTTTAACTGACTTAT 370
QY 1309 ATATATAATTTTTTGTTCATATTGATGTGTCTAGGACGACCTGTGCCAAGTTCT 1368
DB 369 ATATATAATTTTTTGTTCATATTGATGTGTCTAGGACGACCTGTGCCAAGTTCT 310
QY 1369 TAGTGTGTGATGTCGTGGTGTAGACGTAGAAAAGGAATGCAATCCAGACGGT 1428
DB 309 TAGTGTGTGATGTCGTGGTGTAGACGTAGAAAAGGAATGCAATCCAGACGGT 250
QY 1429 TAGTGAATCAGTAAAGCTAGAAATGATCCCGAGCTGTTTATGATAGATAATCTCTCA 1488
DB 249 TAGTGAATCAGTAAAGCTAGAAATGATCCCGAGCTGTTTATGATAGATAATCTCTCA 190
QY 1489 TTCCCGTGGAGAGTTTTCCTGTTCTTTAAAGAGTGATTTGCTGTAGAAAGATGGCACTTA 1548
DB 189 TTCCCGTGGAGAGTTTTCCTGTTCTTTAAAGAGTGATTTGCTGTAGAAAGATGGCACTTA 130
QY 1549 TAACCAAAAGCCCAAGTGTATAGAAATCGTGGTTTTCAGTTTCAGAGAGGGGTGAT 1608
DB 129 TAACCAAAAGCCCAAGTGTATAGAAATCGTGGTTTTCAGTTTCAGAGAGGGGTGAT 70
QY 1609 TTCACACACTACAGTACAGTCTGTATTAAGTTGTTAATAAAGTACATGTTAACTT 1668
DB 69 TTCACACACTACAGTACAGTCTGTATTAAGTTGTTAATAAAGTACATGTTAACTT 10
QY 1669 AAAAAAAAAA 1677
DB 9 AAAAAAAAAA 1

```

RESULT 7
AA634211/c 693 bp mRNA EST 06-MAR-1998
LOCUS AA634211/c
DEFINITION ac72b06.s1 StrataGene Lung (#937210) Homo sapiens CDNA clone
IMAGE:868115 3' similar to gb:L06797 PROBABLE G PROTEIN-COUPLED

ACCESSION AA634211
NID 02557425
VERSION AA634211.1 GI:2557425
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 693)
AUTHORS Haller, J., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kitzman, D., Kuaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393895.

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 835 Std Error: 0.00
Seq primer: -40ml3 fwd. RT from Amersham
High quality sequence stop: 393.
Location/Qualifiers

FEATURES

source

1.693
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:868115"
/clone_1id="Stratagene Lung (#937210)"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. normal lung. Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTGTGTTTGTGTTT 3'"

BASE COUNT 239 a 141 c 124 g 188 t 1 others
ORIGIN

Query Match 28.0%; Score 486; DB 36; Length 693;
Best Local Similarity 99.8%; Pred. No. 1.7e-200;
Matches 536; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1135 TTTCACCTCCAGTACACAGATGTAAAGACTTTTATACGATAAATACTTTT 1194
DB 537 TTTCACCTCCAGTACACAGATGTAAAGACTTTTATACGATAAATACTTTTCT 478
QY 1195 TAGTATACACATTTTTCAGATATAAAGACGTGACCAATATTGACGTTTATGCTTG 1254
DB 477 TAGTATACACATTTTTCAGATATAAAGACGTGACCAATATTGACGTTTATGCTTG 418
QY 1255 TTGGAATTTTGTCTGTGTTCTTTAGTTTGTGAGTTTAATGACTTATTATATA 1314
DB 417 TTGGAATTTTGTCTGTGTTCTTTAGTTTGTGAGTTTAATGACTTATTATATA 358
QY 1315 ATTTTGTGTTTTCATATTGATGTGTCTAGGACGACCTGTGCCAAGTCTTAGTTG 1374
DB 357 ATTTTGTGTTTTCATATTGATGTGTCTAGGACGACCTGTGCCAAGTCTTAGTTG 298
QY 1375 CTGTATGTCGTGTGTAGAGCTGTAGAAAGGAGACTAACATTCCAGAGCGGTAGTGA 1434
DB 297 CTGTATGTCGTGTGTAGAGCTGTAGAAAGGAGACTGAACATTCCAGAGCGGTAGTGA 238
QY 1435 ATCAGCTAAAGCTAGAAATGATCCCGAGCTTTATGCAATAGATAATCTCTCAATCCCG 1494

```


similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
AA781110
NID 92840441
VERSION AA781110.1 GI:2840441
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 792)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1398066.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrp/image/image.html
Insert Length: 1123 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 474.
FEATURES
SOURCE
1..792
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1391178"
/clone_lib="Soares-testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand CDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5',
IGTTACCAATCTGAGTGGAGCGGCCCAATTTTCTTTTCTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 264 a 160 c 154 g 214 t
ORIGIN
Query Match 26.6%; Score 462; DB 38; Length 792;
Best Local Similarity 100.0%; Pred. No. 3.9e-190;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1210 TCAGATATAAAGACTGACCAATATTTGTACAGTTTATTGCTGTGGATTTCCTT 1269
DB 462 TCAGATATAAAGACTGACCAATATTTGTACAGTTTATTGCTGTGGATTTCCTT 403
OY 1270 GTGTTCTTACCTTTTGTGAGAGTTTAATGACTTATTAATTTTGTGTTCA 1329
DB 402 GTGTTCTTACCTTTTGTGAGAGTTTAATGACTTATTAATTTTGTGTTCA 343
OY 1330 TATGATGTGTCTAGCAGGAGCCTGTGGCCAACTTCTAGTGTATGCTGTG 1389
DB 342 TATGATGTGTCTAGCAGGAGCCTGTGGCCAACTTCTAGTGTATGCTGTG 283
OY 1390 TAGGACTGTAGAAAGGAGACTGACATTCAGAGCGGTAGAGATCAGTAACCTAG 1449
DB 282 TAGGACTGTAGAAAGGAGACTGACATTCAGAGCGGTAGAGATCAGTAACCTAG 223

OY 1450 AAATGATCCCGACTGTTTATGCAATATATCTTCACATTCGCCGTGAACGTTTTTCCT 1509
DB 222 AAATGATCCCGACTGTTTATGCAATATATCTTCACATTCGCCGTGAACGTTTTTCCT 163
OY 1510 GTTCTTAGACGATGATTTTCCTGATAGAGATGACCTTATACCAAGCCCAAGTGTA 1569
DB 162 GTTCTTAGACGATGATTTTCCTGATAGAGATGACCTTATACCAAGCCCAAGTGTA 103
OY 1570 TAGAAATGCTGTGTTTTCAGTTTTCAGAGATGGGTGATTCAGACCTACAGTACAG 1629
DB 102 TAGAAATGCTGTGTTTTCAGTTTTCAGAGATGGGTGATTCAGACCTACAGTACAG 43
OY 1630 TCTTGTATTAGTTGTTTAAATAAGTACATGTTAACTTAA 1671
DB 42 TCTTGTATTAGTTGTTTAAATAAGTACATGTTAACTTAA 1
RESULT 12
AA743645/C
LOCUS
DEFINITION
AA743645 556 bp mRNA EST 22-JAN-1998
ny24g09.s1 NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1272736 3'
similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION
NID
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrp/image/image.html
Insert Length: 742 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 389.
FEATURES
SOURCE
1..556
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1272736"
/clone_lib="NCI CGAP GCB1"
/tissue="germinal center B cell"
/lab_host="DH10B"
/note="vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand CDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAGTGGAGCGGCCCAATTTTCTTTTCTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I

VERSION AI344724.1 GI:4081930
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 463)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 14, 1998 this sequence version replaced gi:1797515.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmett-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html
Insert Length: 1208 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 379.
Location/Qualifiers
1.463
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone.lib="IMAGE:1917089"
/tissue.type="NCI CGAP Kid5"
/lab_host="DH10b"
/note="Organ: Kidney; Vector: p77F3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - Oligo(dT) primer [5'
AAGTCGACGATTCGGCGCGCAATATTTTGTTCATATGATGTGTCTAGCAGACCTGT
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p77F3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 171 a 97 c 69 g 126 t
ORIGIN
Query Match 25.3%; Score 439; DB 45; Length 463;
Best Local Similarity 100.0%; Pred. No. 4e-180;
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1238 ACGCTTTTATTCCTGTTGGATTTTGTCTTCTTCTTCTTCTTGAAGTTAA 1297
Db 439 ACGTTTTTATTCCTGTTGGATTTTGTCTTCTTCTTCTTCTTGAAGTTAA 380
OY 1298 TTGACTTATTAATAAATTTTTTTGTTTCATATGATGTGTCTAGCAGACCTGT 1357
Db 379 TTGACTTATTAATAAATTTTTTTGTTTCATATGATGTGTCTAGCAGACCTGT 320
OY 1358 GGGCAGGTTCTAGTTCGTATGCTCTCGTGGTAGGACTATAGAAAAGGACATGAACAT 1417
Db 319 GGGCAGGTTCTAGTTCGTATGCTCTCGTGGTAGGACTATAGAAAAGGACATGAACAT 260
OY 1418 TCCAGAGCGTGTAGTACATCAGTAAAGCTAGAAATGATCCCGAGCTTATGATCATA 1477
Db 259 TCCAGAGCGTGTAGTACATCAGTAAAGCTAGAAATGATCCCGAGCTTATGATCATA 200
OY 1478 TAATCTCTCCATTCGGTGGACGTTTTTCCTTCTTAAGACGTGATTTGGCTGATAGA 1537
Db 199 TAATCTCTCCATTCGGTGGACGTTTTTCCTTCTTAAGACGTGATTTGGCTGATAGA 140

OY 1538 GATGCACCTTATACCAAGGCCAAGTGTATGAATGCTGTTTTCAGTTTCAGG 1597
Db 139 GATGCACCTTATACCAAGGCCAAGTGTATGAATGCTGTTTTCAGTTTCAGG 80
OY 1598 AGTGGTTGATTTCAGCAGCTACAGTGTACAGTCTTATTAAGTTTATAAAGTAC 1657
Db 79 AGTGGTTGATTTCAGCAGCTACAGTGTACAGTCTTATTAAGTTTATAAAGTAC 20
OY 1658 ATGTTAACTTAAAAAAA 1676
Db 19 ATGTTAACTTAAAAAAA 1
RESULT 15
AI002547/c 482 bp mRNA EST 10-NOV-1998
LOCUS oq90e01.s1 NCI-CGAP_Kid6 Homo sapiens cDNA clone IMAGE:1593624 3'
DEFINITION similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AI002547
NID g3202881
VERSION AI002547.1 GI:3202881
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 482)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2284974.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmett-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html
Insert Length: 1054 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 393.
Location/Qualifiers
1.482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="964E2; 14924.3"
/clone.lib="IMAGE:1593624"
/clone.lib="NCI CGAP Kid6"
/sex="mixed"
/tissue.type="Kidney tumor"
/lab_host="GOLR (kanamycin resistant)"
/note="Organ: kidney; Vector: Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5'
GAATTCGGACGAG 3' 3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.0 kb."

BASE COUNT 179 a 99 c 72 g 132 t
ORIGIN
Query Match 25.2%; Score 437; DB 41; Length 482;
Best Local Similarity 100.0%; Pred. No. 2.9e-179;
Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1239 CAGTTTTTATTCCTGTTGGATTTTGTCTTCTTCTTCTTCTTGAAGTTAA 1298

```

|||||
Db 437 CAGTTTATGCTGTGATTTGCTGTGTTCTTATGTTTGTGAAGTTAAT 378
Oy 1299 TGACTATTTATTAATTTTTTTTGTTCATATGATGTGTCTAGGAGACCTGTG 1358
Db 377 TGACTATTTATTAATTTTTTTTGTTCATATGATGTGTCTAGGAGACCTGTG 318
Oy 1359 GGCAGTTCTAGTCTGTGTATGTCGTGTGTAGACTGTAGAAAAGGGAACATG 1418
Db 317 GCCAAGTCTTACTTGTCTGTATGTCTGTGTAGACTGTAGAAAAGGGAACATG 258
Oy 1419 CCAGAGCGTGTAGTGAATCACTGAATGAATGATCCCGAGCTGTTATGATATGAT 1478
Db 257 CCAGAGCGTGTAGTGAATCACTGAATGAATGATCCCGAGCTGTTATGATATGAT 198
Oy 1479 AATCTCTCATTTCCCGTGGAGACGTTTTTCTGTCTTAAACGATGTTTGTCTGTAAG 1538
Db 197 AATCTCTCATTTCCCGTGGAGACGTTTTTCTGTCTTAAACGATGTTTGTCTGTAAG 138
Oy 1539 ATGGCACTTTAACCAGCCCAAGGATGATGAATGCTGTTTTCAGTTTTCAGGA 1598
Db 137 ATGGCACTTTAACCAGCCCAAGGATGATGAATGCTGTTTTCAGTTTTCAGGA 78
Oy 1599 GTGGGTTGATTGACGACCTACAGTGTACAGTCTGTATTAAGTTGTTAATAAAGTACA 1658
Db 77 GTGGGTTGATTGACGACCTACAGTGTACAGTCTGTATTAAGTTGTTAATAAAGTACA 18
Oy 1659 TGTAAACTTAAAAAA 1675
Db 17 TGTAAACTTAAAAAA 1

RESULT 16
AI249788/c 524 bp mRNA EST 03-FEB-1999
LOCUS qx51c10.x1 NCI-CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2004862 3'
DEFINITION similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AI249788
NID g3846317
VERSION AI249788.1 GI:3846317
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 524)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TUMOR National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Aug 21, 1998 this sequence version replaced.

FEATURES
source
1..524
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2004862"
/clone_lib="NCI-CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/notes="Organ: pancreas; Vector: pCMV-Sport6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"
BASE COUNT 195 a 103 c 73 g 151 t 2 others

```

```

ORIGIN
Query Match 25 0%; Score 435; DB 44; Length 524;
Best Local Similarity 100.0%; Pred. No. 2.1e-178;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1246 TATGCTTGTGGATTTTGTCTGTTGTTCTTACTTTTGTGAAGTTAATGACTTA 1305
Db 435 TATGCTTGTGGATTTTGTCTGTTGTTCTTACTTTTGTGAAGTTAATGACTTA 376
Oy 1306 TTTATATTAATTTTTTTTGTTCATATTTGATGTGTCTAGGAGACCTGTGGCAAGT 1365
Db 375 TTTATATTAATTTTTTTTGTTCATATTTGATGTGTCTAGGAGACCTGTGGCAAGT 316
Oy 1366 TCTTATGCTGTATGTCTGTGTAGACTGTAGAAAAGGGAACATGCCAGAGC 1425
Db 315 TCTTATGCTGTATGTCTGTGTAGACTGTAGAAAAGGGAACATGCCAGAGC 256
Oy 1426 GTGTAGTAATCACTGAAGCTGAATGATCCCGAGCTGTTATGATATATCTCT 1485
Db 255 GTGTAGTAATCACTGAAGCTGAATGATCCCGAGCTGTTATGATATATCTCT 196
Oy 1486 CCAATCCCGTGAAGCTTTTCTGTCTTAAAGAGTATTTTGTCTGTAAGAGTGGCAG 1545
Db 195 CCAATCCCGTGAAGCTTTTCTGTCTTAAAGAGTATTTTGTCTGTAAGAGTGGCAG 136
Oy 1546 TTTAATCCAAAGCCCAAGGATGATGAATGCTGTTTTCAGTTTTCAGAGTGGGTT 1605
Db 135 TTTAATCCAAAGCCCAAGGATGATGAATGCTGTTTTCAGTTTTCAGAGTGGGTT 76
Oy 1606 GATTTCAGACCTACAGTGTACAGTCTGTATTAAGTTGTTAATAAGTACTGTTAA 1665
Db 75 GATTTCAGACCTACAGTGTACAGTCTGTATTAAGTTGTTAATAAGTACTGTTAA 16
Oy 1666 CTTAAAAAATAAAAAA 1680
Db 15 CTTAAAAAATAAAAAA 1

RESULT 17
AA259199/c 428 bp mRNA EST 20-AUG-1997
LOCUS nc17b08.r1 NCI-CGAP_Pri1 Homo sapiens cDNA clone IMAGE:1008375
DEFINITION similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA259199
NID g1894641
VERSION AA259199.1 GI:1894641
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 428)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TUMOR National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1394051.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,
M.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```

Insert length: 879 Std Error: 0.00
 Seq primer: 28m13 rev1 ET from Amersham
 High quality sequence stop: 330.
 Location/Qualifiers
 1. 428

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1008375"
 /clone_1lb="NCI_CGAP_Pp1"
 /sex="Male"
 /dev_stage="45 years old"
 /lab_host="DH10B"
 /note="Vector: PAMp10: Site.1: Not1: Site.2: EcoRI: 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected, histologically normal prostate epithelial cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into PAMp10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

BASE COUNT 165 a 94 c 62 g 107 t
 ORIGIN

Query Match 24.6%; Score 428; DB 30; Length 428;
 Best Local Similarity 100.0%; Pred. No. 2.4e-175;

Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1241 GTTTATGCTGCTGGATTTTGTCTGTGTTCTTATGTTTGTGAAGTTAATG 1300
 DB 428 GTTTATGCTGCTGGATTTTGTCTGTGTTCTTATGTTTGTGAAGTTAATG 369
 QY 1301 ACTATATTAATAATTTTTTGTGTTTCAATATGATGTGTCTAGCAGACCTGTGC 1360
 DB 368 ACTATATTAATAATTTTTTGTGTTTCAATATGATGTGTCTAGCAGACCTGTGC 309
 QY 1361 CAGTCTAGTGTGTGATGCTGCTGTAGAGCTGTAGAAAAGGAATGAACATTC 1420
 DB 308 CAGTCTAGTGTGTGATGCTGCTGTAGAGCTGTAGAAAAGGAATGAACATTC 249
 QY 1421 AGAGCGTGTAGTAATCAGGTAAAGTATGATCCAGCTGTTATGATAGATAA 1480
 DB 248 AGAGCGTGTAGTAATCAGGTAAAGTATGATCCAGCTGTTATGATAGATAA 189
 QY 1481 TCTCTCATTTCCCGTGAACGTTTTCTCTGTTTAAAGCGTATTTGCTGTAGAAGAT 1540
 DB 188 TCTCTCATTTCCCGTGAACGTTTTCTCTGTTTAAAGCGTATTTGCTGTAGAAGAT 129
 QY 1541 GGCACCTATACCAAGCCCAAGGTGTATAGAAATGCTGTTTTCAGTTTCAGAGAT 1600
 DB 128 GGCACCTATACCAAGCCCAAGGTGTATAGAAATGCTGTTTTCAGTTTCAGAGAT 69
 QY 1601 GGGTGTATTTACGACCTACAGTGTACAGTCTTGTATTAAGTTGTTAATAAAGTACATG 1660
 DB 68 GGGTGTATTTACGACCTACAGTGTACAGTCTTGTATTAAGTTGTTAATAAAGTACATG 9
 QY 1661 TTAACCTT 1668
 DB 8 TTAACCTT 1

RESULT 18
 LOCUS AI300764 473 bp mRNA EST 01-FEB-1999
 DEFINITION qn50c09.x1 NCI_CGAP_Kids Homo sapiens cDNA clone IMAGE:1901680 3'
 similar to gb:105797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
 HOMOLOG (HUMAN); mRNA sequence.
 ACCESSION AI300764
 NID q3960110
 VERSION AI300764.1 GI:3960110

KEYWORDS

EST.
 human.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 473)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

On Aug 21, 1998 this sequence version replaced.

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLN at:
www.bio.lnl.gov/bbrp/image/image.html

Insert length: 995 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 365.
 Location/Qualifiers

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="3p14-p21"
 /clone="IMAGE:1901680"
 /clone_1lb="NCI_CGAP_Kids"
 /tissue_type="2 pooled tumors (clear cell type)"
 /lab_host="DH10B"
 /note="Organ: Kidney; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' AACTGAGAAATTCGGGGCCGCAATATTTTGTGTTTGTGTTT 3'), double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bernaldo. "

BASE COUNT 175 a 98 c 70 g 128 t 2 others
 ORIGIN

Query Match 24.6%; Score 427; DB 44; Length 473;
 Best Local Similarity 100.0%; Pred. No. 5.3e-175;

Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1246 TATTCCTGTGATTTTGTCTGTTTCTTATGTTTGTGAAGTTAATGACTTA 1305
 DB 427 TATTCCTGTGATTTTGTCTGTTTCTTATGTTTGTGAAGTTAATGACTTA 368
 QY 1306 TTAATAATAATTTTTTGTGTTTCAATATGATGTGTCTAGAGGAGACCTGGGCAAGT 1365
 DB 367 TTAATAATAATTTTTTGTGTTTCAATATGATGTGTCTAGAGGAGACCTGGGCAAGT 308
 QY 1366 TCTTAGTGTGATGCTCTGCTGTAGAGCTGTAGAAAAGGAGCAATCCAGAGC 1425
 DB 307 TCTTAGTGTGATGCTCTGCTGTAGAGCTGTAGAAAAGGAGCAATCCAGAGC 248
 QY 1426 GTTAGTGAATCAGTAAGCTAGAAATGATCCCGAGCTGTTATATCATAGATATCTCT 1485
 DB 247 GTTAGTGAATCAGTAAGCTAGAAATGATCCCGAGCTGTTATATCATAGATATCTCT 188
 QY 1486 CCATCCCGTGAAGCTTTTCTGTTCTTAAGAGCTGATTTTGCTGTGAAGATGGCAC 1545
 DB 187 CCATCCCGTGAAGCTTTTCTGTTCTTAAGAGCTGATTTTGCTGTGAAGATGGCAC 128

QY 1546 TTATACCAAGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTT 1605
 |||||||
 Db 127 TTATACCAAGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTT 68
 |||||||
 QY 1606 GATTTCGACCTACAGTGTACAGTCTTGTATTAGTTGTTAAATAAGTACAGTTAA 1665
 |||||||
 Db 67 GATTTCGACCTACAGTGTACAGTCTTGTATTAGTTGTTAAATAAGTACAGTTAA 8
 |||||||
 QY 1666 CTTAAAA 1672
 |||||||
 Db 7 CTTAAAA 1
 |||||||
 RESULT 19
 A1022030/c 535 bp mRNA EST 28-AUG-1998
 LOCUS A1022030/1
 DEFINITION ow72a11.x1 Soares_fetal_liver_spleen_INFLS.S1 Homo sapiens CDNA
 clone IMAGE:1652348 3' similar to gb:L06797 PROBABLE G
 PROTEIN-COUPLED RECEPTOR LCRI HOMOLOG (HUMAN); mRNA sequence.
 A1022030
 A1022030
 A1022030.1 GI:3239383
 VERSION 93239383
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 535)
 NC1-CCAP http://www.ncbi.nlm.nih.gov/nc1ccap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Jan 9, 1998 this sequence version replaced gi:785673.
 CONTACT: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 975 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 189.
 Location/Qualifiers
 1..535
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1652348"
 /clone_1fb="Soares_fetal_liver_spleen_INFLS.S1"
 /sex="male"
 /dev_stage="20 week-post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
 with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
 This is a subtracted version of the original Soares fetal
 liver spleen INFLS library. 1st strand cDNA was primed
 with a Pac I - oligo(dT) primer [5',
 AACTGGAAGATTAATTAAGATCTTTTCTTTTCTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 205 a 108 c 79 g 142 t 1 others
 ORIGIN

Query Match 24.4% Score 423 DB 41 Length 535;
 Best Local Similarity 99.8% Pred. No. 3.3e-173;
 Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1195 TAAGTTACACATTTTTCAGATATAAAGACGACCAATATTGTACAGTTTATGCTTG 1254
 |||||||
 Db 474 TAAGTTACACATTTTTCAGATATAAAGACGACCAATATTGTACAGTTTATGCTTG 415

QY 1255 TTGATTTTTCCTGCTGCTTCTTTAGTTTGTGGAAGTTAATGACTATTATATTA 1314
 |||||||
 Db 414 TTGATTTTTCCTGCTGCTTCTTTAGTTTGTGGAAGTTAATGACTATTATATTA 355
 |||||||
 QY 1315 ATTTTTCCTGCTGCTTCTTTAGTTTGTGGAAGTTAATGACTATTATATTA 1374
 |||||||
 Db 354 ATTTTTCCTGCTGCTTCTTTAGTTTGTGGAAGTTAATGACTATTATATTA 295
 |||||||
 QY 1375 CTGATGTCCTGCTGCTTCTTTAGTTTGTGGAAGTTAATGACTATTATATTA 1434
 |||||||
 Db 294 CTGATGTCCTGCTGCTTCTTTAGTTTGTGGAAGTTAATGACTATTATATTA 235
 |||||||
 QY 1435 ATCAGCTAAAGCTAGATATATATATATATATATATATATATATATATATAT 1494
 |||||||
 Db 234 ATCAGCTAAAGCTAGATATATATATATATATATATATATATATATATATAT 175
 |||||||
 QY 1495 TGAAGCTTTTCTGCTGCTTCTTTAGTTTGTGGAAGTTAATGACTATTATAT 1554
 |||||||
 Db 174 TGAAGCTTTTCTGCTGCTTCTTTAGTTTGTGGAAGTTAATGACTATTATAT 115
 |||||||
 QY 1555 AAGCCCAAGCTAGATATATATATATATATATATATATATATATATATAT 1614
 |||||||
 Db 114 AAGCCCAAGCTAGATATATATATATATATATATATATATATATATATAT 55
 |||||||
 QY 1615 ACCACAGTGTACAGTCTTGTATTAGTTGTTAAATAAGTACATGTTAACTT 1668
 |||||||
 Db 54 ACCACAGTGTACAGTCTTGTATTAGTTGTTAAATAAGTACATGTTAACTT 1

RESULT 20
 AAI48292/c 583 bp mRNA EST 06-AUG-1997
 LOCUS AAI48292/1
 DEFINITION z045e12.s1 Stragane endothelial cell 937223 Homo sapiens CDNA
 clone IMAGE:589870 3' similar to gb:L06797 PROBABLE G
 PROTEIN-COUPLED RECEPTOR LCRI HOMOLOG (HUMAN); mRNA sequence.
 AAI48292
 AAI48292
 AAI48292.1 GI:1717715
 VERSION 91717715
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 583)
 Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
 Chissee,S., Dietrich,N., Dubuque,T., Favellio,A., Gish,W.,
 Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
 Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
 Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierly-Meg,J.,
 Trevasakis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
 and Marra,M.
 Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-828 (1996)
 MEDLINE 97044478
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1394359.

TITLE Contact: Wilson RK
 JOURNAL Washington University School of Medicine
 MEDLINE 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 FAX: 314 286 1800
 Tel: 314 286 1800
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 888 Std Error: 0.00
 Seq primer: -40m13 fwd. from Amersham
 High quality sequence stop: 386.
 Location/Qualifiers
 1..583
 /organism="Homo sapiens"
 /db_xref="GDB:4621256"
 /db_xref="taxon:9606"

FEATURES
 SOURCE

Db 349 TTATATATGATGTCGTCTGAGCAGACCTGTGCGCAAGTTCTTAGTCTGATGTCI 290

QY 1385 CGTGTAGAGACTGTAGAAAAAGGAAGTGAACATTCACAGAGCGGTAGTGAATCAGTAAA 1444
 |||
 Db 289 CGTGTAGAGACTGTAGAAAAAGGAAGTGAACATTCACAGAGCGGTAGTGAATCAGTAAA 230
 |||

QY 1445 GCTAAGAAATGATCCCGAGCTTTTATGATAGATATATCTCCATCTCCGTTGGAACGTTT 1504
 |||
 Db 229 GCTAAGAAATGATCCCGAGCTTTTATGATAGATATATCTCCATCTCCGTTGGAACGTTT 170
 |||

QY 1505 TTCTGTCTTCTTAAGACGATTTGCTGTAGAGATGACATTTATACCAAGCCCAAG 1564
 |||
 Db 159 TTCTGTCTTCTTAAGACGATTTGCTGTAGAGATGACATTTATACCAAGCCCAAG 110
 |||

QY 1565 TGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTGAATTCAGACCTACAGCTG 1624
 |||
 Db 109 TGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTGAATTCAGACCTACAGCTG 50
 |||

QY 1525 TACAGTCTTGTATTAAGTTTATATAAGTACATGTAACCTTAAAAA 1673
 |||
 Db 49 TACAGTCTTGTATTAAGTTTATATAAGTACATGTAACCTTAAAAA 1

RESULT 24
 AA750743/c 509 bp mRNA EST 16-FEB-1998
 LOCUS n210809.s1 NCI-CGAP_GCB1 Homo sapiens CDNA clone IMAGE:1287353 3'
 DEFINITION similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
 HOMOLOG (HUMAN); mRNA sequence.
 AA750743
 92809673
 AA750743.1 GI:2809673
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 509)
 REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1397905.

FEATURES
 source
 Insert length: 739 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 340.
 Location/Qualifiers
 1. 509
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1287353"
 /clone_1b="NCI-CGAP_GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"
 /note="Vector: p713D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells selected for
 germinal center B cells by flow sorting (CD20+, IgD+),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman

(NCI) and Dr. Gerald Marti (CEBR). CDNA synthesis was
 primed with a Not I - 0.190(dt) primer
 [5'-TGTTCACATCTGAGTGGAGCGCCGCTCATTTTCTTTTCTTTT-
 3']. Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified p713 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 188 a 99 c 73 g 149 t
 ORIGIN

Query Match 23.2%; Score 403; DB 38; Length 509;
 Best Local Similarity 100.0%; Pred. No. 1.5e-164;
 Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1279 TAGTTTGTGAAGTTTATGACTTATTTATATAATTTTGTTCATATGATGT 1338
 |||
 Db 403 TAGTTTGTGAAGTTTATGACTTATTTATATAATTTTGTTCATATGATGT 344
 |||

QY 1339 GTGCTAGCAGACCTGTGCGCAAGTTCTTAGTCTGTAATGCTGTAGACGTG 1398
 |||
 Db 343 GTGCTAGCAGACCTGTGCGCAAGTTCTTAGTCTGTAATGCTGTAGACGTG 284
 |||

QY 1399 AGAAAGGGAAGTGAACATTCACAGCGCTAGTGAATCAGCTAAGTAAATGATCC 1458
 |||
 Db 283 AGAAAGGGAAGTGAACATTCACAGCGCTAGTGAATCAGCTAAGTAAATGATCC 224
 |||

QY 1459 CCAGCTGTTTATGATAGATATCTCTCCATCTCCGTTGAGACGTTTCTGTTCTTAAG 1518
 |||
 Db 223 CCAGCTGTTTATGATAGATATCTCTCCATCTCCGTTGAGACGTTTCTGTTCTTAAG 164
 |||

QY 1519 ACGTATTTTCTGTAGAAAGTGCACCTATTAACCAAGCCCAAGTGTATAGAAATGC 1578
 |||
 Db 163 ACGTATTTTCTGTAGAAAGTGCACCTATTAACCAAGCCCAAGTGTATAGAAATGC 104
 |||

QY 1579 TGGTTTTTCAGTTTTCAGAGTGGTGGTGAATTCAGACCTACAGTGTACGTTGAT 1638
 |||
 Db 103 TGGTTTTTCAGTTTTCAGAGTGGTGGTGAATTCAGACCTACAGTGTACGTTGAT 44
 |||

QY 1639 AAGTGTATATAAGTACATGTTAACTTAAAAA 1681
 |||
 Db 43 AAGTGTATATAAGTACATGTTAACTTAAAAA 1

RESULT 25
 A1358446/c 417 bp mRNA EST 15-FEB-1999
 LOCUS A1358446
 DEFINITION qx19c11.x1 NCI-CGAP_Lym12 Homo sapiens CDNA clone IMAGE:2001812 3'
 similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
 HOMOLOG (HUMAN); mRNA sequence.
 A1358446
 94110067
 A1358446.1 GI:4110067
 EST.
 VERSION A1358446.1
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 417)
 REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Aug 21, 1998 this sequence version replaced.

FEATURES
 source
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Life Technologies catalog #: 11547-015
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.lnl.gov/bb-p/image/image.html

Insert Length: 899 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 321.
Location/Qualifiers

FEATURES

Source

1. 417
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2001812"
/clone_lib="NCI CGAP_Lym12"
/tissue_type="lymphoma, follicular mixed small and large cell"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: pCMV-SPORT6; Site: 1: SalI; Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.25 kb. Life Technologies catalog #: 11547-015"
BASE COUNT 148 a 88 c 61 g 120 t
ORIGIN

Query Match 23.2%; Score 403; DB 45; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.6e-164;
Matches 403; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1280 AGTTTGTGAAGTTAATGACTATTATATAATTTTGTTCATATGATGTC 1339
DB 403 ACTTTTGTGAAGTTAATGACTATTATATAATTTTGTTCATATGATGTC 344
QY 1340 TGTCTAGGAGACCGTGGCCAGTTCTAGTGTCTGCTGCTGCTAGACTGTA 1399
DB 343 TGTCTAGGAGACCGTGGCCAGTTCTAGTGTCTGCTGCTGCTGCTAGACTGTA 284
QY 1400 GAAAGGAGACTGACATTCACAGAGCGTAGTGAATCAGTAAGCTAGAAATGCC 1459
DB 283 GAAAGGAGACTGACATTCACAGAGCGTAGTGAATCAGTAAGCTAGAAATGCC 224
QY 1460 CAGCTGTTATGATGATATATCTCTCCATTCCTCCGGAACGTTTCCGTTCTTAAGA 1519
DB 223 CAGCTGTTATGATGATATATCTCTCCATTCCTCCGGAACGTTTCCGTTCTTAAGA 164
QY 1520 CGGATTTGCTGTAGAAGTGGCAGCTTATACCAAGGAGTGTAGAAATGCT 1579
DB 163 CGGATTTGCTGTAGAAGTGGCAGCTTATACCAAGGAGTGTAGAAATGCT 104
QY 1580 GGTTTTTCAGTTTTCAGGAGTGGTGTGATTTTCAGCAGCTACAGTCTTGATTA 1639
DB 103 GGTTTTTCAGTTTTCAGGAGTGGTGTGATTTTCAGCAGCTACAGTCTTGATTA 44
QY 1640 AGTTGTTAATAAAGTACATGTTAACTTAAAAA 1682
DB 43 AGTTGTTAATAAAGTACATGTTAACTTAAAAA 1

RESULT 26
LOCUS A1660355 401 bp mRNA EST 10-MAY-1999
DEFINITION we62b07.x1 Soares_thymus_NHEFth Homo sapiens cDNA clone
IMAGE:2345653 3' similar to gb:106797 PROBABLE G PROTEIN-COUPLED
RECEPTOR LCRI HOMOLOG (HUMAN); mRNA sequence.
ACCESSION A1660355
NID 94763925
VERSION A1660355.1 GI:4763925
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 401)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index

JOURNAL COMMENT

Unpublished (1997)
On Mar 10, 1998 this sequence version replaced gi:2949344.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40UP from Gibco.

FEATURES

Source

1. 401
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2345653"
/clone_lib="Soares_thymus_NHEFth"
/dev_stage="fetal"
/lab_host="DH10B (phage-resistant)"
/note="Organ: thymus; pooled; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site: 1: Not I; Site: 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5', TGTTACCATCTGAGTGGAGCGCGCCGACGTTTGTCTTTTGTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 107 a 59 c 82 g 153 t
ORIGIN

Query Match 23.1%; Score 401; DB 49; Length 401;
Best Local Similarity 100.0%; Pred. No. 1.2e-163;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1181 AATAACTTTTATTAAGTACACATTTTTCAGATATATAAAGACGACCAATATGTACA 1240
DB 1 AATAACTTTTATTAAGTACACATTTTTCAGATATATAAAGACGACCAATATGTACA 60
QY 1241 GTTTTATGCTGTGGATTTTGTCTGTGTTCTTGTAGTTTGTGAAGTTAATG 1300
DB 61 GTTTTATGCTGTGGATTTTGTCTGTGTTCTTGTAGTTTGTGAAGTTAATG 120
QY 1301 ACTTATTATTAATTTTGTGTTTCATATGATGTGTCTAGCAGACCTGTGCG 1360
DB 121 ACTTATTATTAATTTTGTGTTTCATATGATGTGTCTAGCAGACCTGTGCG 180
QY 1361 CAAGTCTTAGTTCGTGTATGTCTGTGTAGACCTGTAGAAAAGGAGACTGAACATTCC 1420
DB 181 CAAGTCTTAGTTCGTGTATGTCTGTGTAGACCTGTAGAAAAGGAGACTGAACATTCC 240
QY 1421 AGAGCGTAGTGAATCAGTAAAGCTAGAAATGATCCAGCTGTTATGCTATAGATAA 1480
DB 241 AGAGCGTAGTGAATCAGTAAAGCTAGAAATGATCCAGCTGTTATGCTATAGATAA 300
QY 1481 TCTCTCCATTCGCCGTGAACGTTTTCCTGTTCTTAAGACGATTTTGTGTAGAAAT 1540
DB 301 TCTCTCCATTCGCCGTGAACGTTTTCCTGTTCTTAAGACGATTTTGTGTAGAAAT 360
QY 1541 GGCACCTTATACCAAGCCCAAGTGTATAGAAATGCTGG 1581
DB 361 GGCACCTTATACCAAGCCCAAGTGTATAGAAATGCTGG 401

RESULT 27
LOCUS AA386000/c 393 bp mRNA EST 21-APR-1997
DEFINITION EST9758 Pancreas tumor I Homo sapiens cDNA 3' end similar to
guanine nucleotide-binding protein HM89, mRNA sequence.
ACCESSION AA386000
NID 52038337
VERSION AA386000.1 GI:2038337
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
AUTHORS Adams, M.D., Kertlavage, A.R., Fleischmann, R.D., Fulton, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, Wai, C., Clifton, R.A., Cline, T.R., Cotton, M.D., Earle, Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nauyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dime, D., Feng, D.-F., Ferrite, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M., and Venter, J.C.
TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl.), 3-174 (1995)
MEDLINE 95026280
COMMENT On Dec 30, 1996 this sequence version replaced gi:1530936.
Other ESTs: EST9739 T0C16975
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699055
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13-21.

FEATURES
SOURCE location/Qualifiers
1..393
/organism="Homo sapiens"
/db_xref="ATCC (Inhost):190162"
/db_xref="taxon:9606"
/clone_lib="Pancreas tumor I"
/dev_stage="adult"
/note="Organ: pancreas; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 142 a 86 c 60 g 105 t
ORIGIN

Query Match 22.6%; Score 393; DB 33; Length 393;
Best local Similarity 100.0%; Pred. No. 3.4e-160;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1276 CTTTGGTTTGTGAAGTTAATGACTATTATTAATTTTGTTCATATTGA 1335
|||||
Db 393 CTTTGGTTTGTGAAGTTAATGACTATTATTAATTTTGTTCATATTGA 334
Oy 1336 TGTGTGTAGGACAGACTGTGGCCAAAGTTTGTGCTGTATGCTGTAGGAC 1395
|||||
Db 333 TGTGTGTAGGACAGACTGTGGCCAAAGTTTGTGCTGTATGCTGTAGGAC 274
Oy 1396 TGTAAAAAGGAAGTGAACATCCAGAGCGTGTAGTGAATCAGTAAAGCTGAAGAATGA 1455
|||||
Db 273 TGTAAAAAGGAAGTGAACATCCAGAGCGTGTAGTGAATCAGTAAAGCTGAAGAATGA 214
Oy 1456 TCCCGAGCTGTATGATAGATATCTCATTCCTCCGTGAAGCTTTTCTGTCTT 1515
|||||
Db 213 TCCCGAGCTGTATGATAGATATCTCATTCCTCCGTGAAGCTTTTCTGTCTT 154
Oy 1516 AAGACGTGATTTTGTCTGTGAAGATGGCACTTAATACCAAGCCCAAGTGTATGAAA 1575

Db 153 AAGACGTGATTTTGTCTGTAGACATGGCACTTATTAACCAAGCCCAAGGTATGAAA 94
Oy 1576 TCGTGGTTTTCAGTTTTCAGAGAGTGGGTGATTTTCAGACCTTACAGTGTGTTGT 1635
|||||
Db 93 TCGTGGTTTTCAGTTTTCAGAGAGTGGGTGATTTTCAGACCTTACAGTGTGTTGT 34
Oy 1636 ATTAAGTTGTTAATAAAGTACATGTTAACTT 1668
|||||
Db 33 ATTAAGTTGTTAATAAAGTACATGTTAACTT 1

RESULT 28
AA479467 521 bp mRNA EST 08-AUG-1997
LOCUS zvl7110.r1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:753931 5'
DEFINITION similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA479467
NID G2208023
VERSION AA479467.1 GI:2208023
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theisling, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393144.

FEATURES
SOURCE location/Qualifiers
1..521
/organism="Homo sapiens"
/db_xref="GDB:5976803"
/db_xref="taxon:9606"
/map="956G09:1"
/clone="IMAGE:753931"
/clone_lib="Soares_NHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of 2NBHM, pregnant uterus NBHM, and fetal heart NBHM19W were mixed, and ss circles were made in vitro. Following NRP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
BASE COUNT 132 a 104 c 103 g 182 t
ORIGIN

Query Match 22.2%; Score 386; DB 34; Length 521;
Best local Similarity 99.6%; Pred. No. 3.4e-157;

	Matches	486	Conservative	0:	Mismatches	2:	Indels	0:	Gaps	0:
OY	890	TGCAATCATCAGCAGGGGTGTGAGTTTGACACACTGTGCACAAGTGGATTCCATCA	949							
Db	1	TGGAATATCATCAACCAGGGTGTGAGTTTGAAACACTGTGCACAAGTGGATTTCATCA	60							
OY	950	CCGAGGCCCTGAGCTTCTTCCACTTTGTCTCGAACCCCATCCTCTAATGCTTCTCTGAG	1009							
Db	61	CCGAGGCCCTTAGCTTCTTCCACTTTGTCTCGAACCCCATCCTCTAATGCTTCTCTGAG	120							
OY	1010	CCAATTTTAAAACCTCTGCCAGCACGCACTCACTCTGTGACGAGGGSTCCAGCCTTA	1069							
Db	121	CCAATTTTAAAACCTCTGCCAGCACGCACTCACTCTGTGACGAGGGSTCCAGCCTTA	180							
OY	1070	AGATCCTCTCCAAGGAAGAAGGAGGTGGACATTCATCCTTTCCATGAGTGCAGCTTT	1129							
Db	181	AGATCCTCTCCAAGGAAGAAGGAGGTGGACATTCATCCTTTCCATGAGTGCAGCTTT	240							
OY	1130	CAGTTTTCTACCTCCAGCTACACAGATGTAAAAAGCTTTTTTTATAGATAAATACTT	1189							
Db	241	CAGTTTTCTACCTCCAGCTACACAGATGTAAAAAGCTTTTTTTATAGATAAATACTT	300							
OY	1190	TTTTTTAAGTTACACATTTTTCAGATATATAAGACTGACCAATATGTACAGTTTATT	1249							
Db	301	TTTTTTAAGTTACACATTTTTCAGATATATAAGACTGACCAATATGTACAGTTTATT	360							
OY	1250	GCTGTGTGATTTTGTCTGTGTGTTCTTTTACCTTTTGTGAATTTAATGACTTATTA	1309							
Db	361	GCTGTGTGATTTTGTCTGTGTGTTCTTTTACCTTTTGTGAATTTAATGACTTATTA	420							
OY	1310	TATAAATTTTTTTTGTTCATATTTGATGTGTCTCTAGGACGACCGTGGCAAGTTCTT	1369							
Db	421	TATAAATTTTTTTTGTTCATATTTGATGTGTCTCTAGGACGACCGTGGCAAGTTCTT	480							
OY	1370	AGTTGCTG 1377								
Db	481	AGTTGCTG 488								
RESULT 29	A1159856/c									
LOCUS	A1159856	504 bp	mRNA	EST	26-OCT-1998					
DEFINITION	gb15c09.x1 NC1 CGAP Brn2 Homo sapiens cDNA clone IMAGE:1703632	3'								
	similar to gb106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI									
ACCESSION	A1159856									
NID	93693236									
VERSION	A1159856.1	GI:3693236								
KEYWORDS	EST.									
SOURCE	human.									
ORGANISM	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;									
	Eutheria; Primates; Catarrhini; Hominiidae; Homo.									
REFERENCE	1 (bases 1 to 504)									
AUTHORS	NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.									
TITLE	National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGP), Tumor Gene Index									
JOURNAL	Unpublished (1998)									
COMMENT	On Jan 19, 1998 this sequence version replaced gi:2282071.									
	Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.L.Strausberg@nih.gov Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D... Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can b found through the I.M.A.G.E. Consortium/LINL at: www-bio.lnl.gov/db/rp/image/image.html									

FEATURES					
source	Insert Length: 623 Std Error: 0.00 Seq primer: -40ml3 fwd. RT from Amersham High quality sequence stop: 319. Location/Qualifiers				
	1..504				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:1703632"				
	/clone_1bp="NCI CGAP_Brn23"				
	/russue_type="gliblastoma (pooled)"				
	/lab_host="DH10B"				
	/note="Organ: brain; Vector: pT7R3D-Pac (Pharmacia) with a modified polylinker; Site1: Not I ; Site-2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACCAATCTGAGTCAGCGAGCGCCGCAGATCTTTTTTTTTTTTTTTTTT T3'] ; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and Eco RI sites of the Not I and Eco RI sites of the modified pT7R3 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."				
BASE COUNT	185 a 102 c 71 g 146 t				
ORIGIN					
Query Match	22.2%; Score 386; DB 43; Length 504;				
Best Local Similarity	100.0%; Pred. No. 3.4e-157;				
Matches	386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Oy	1296	AATTACTATTATATAAATTTTTTTTGTTTCATATGATGTGTGCTTAGCAGGACCT	1355		
Db	386	AATTACTATTATATAAATTTTTTTTGTTTCATATGATGTGTCTAGGAGACCT	327		
Oy	1356	GTCGCCAACTTCTAGTTCGTGTATGTCCTCGTGTAGGACCTAGAAAAGGAACCTGAAC	1415		
Db	326	GTCGCCAACTTCTAGTTCGTGTATGTCCTCGTGTAGGACCTAGAAAAGGAACCTGAAC	267		
Oy	1416	ATTCCAGAAGCTGTAGTGAATCAGCTAAAGCTAGAAATATATCCCCAGCTGTTATGCATA	1475		
Db	266	ATTCCAGAAGCTGTAGTGAATCAGCTAAAGCTAGAAATATATCCCCAGCTGTTATGCATA	207		
Oy	1476	GATATCTCTCCATTCOCCTGGAGAGTTTTCCTGTCTTAAGACGTGATTTGGCTGTAG	1535		
Db	206	GATATCTCTCCATTCOCCTGGAGAGTTTTCCTGTCTTAAGACGTGATTTGGCTGTAG	147		
Oy	1536	AAGATGGCACTATATACCACAAGCCCAGAAGTGATAGAAATGCTGTTTTCAGTTTCA	1595		
Db	146	AAGATGGCACTATATACCACAAGCCCAGAAGTGATAGAAATGCTGTTTTCAGTTTCA	87		
Oy	1596	GGAGTGGGTGATTTCAGCACCTACAGTGTAAAGTGTGATTAAGTTGTTATATAAAGT	1655		
Db	86	GGAGTGGGTGATTTCAGCACCTACAGTGTAAAGTGTGATTAAGTTGTTATATAAAGT	27		
Oy	1656	ACATGTTAACTTAAAAAIAAAAAAAAA 1681			
Db	26	ACATGTTAACTTAAAAAIAAAAAAAAA 1			
RESULT 30					
AI193140/C					
LOCUS	AI193140	475 bp	mRNA	EST	29-OCT-1998
DEFINITION	qe40f12.s1 Soares-fetal.Lung-NBH119W Homo sapiens cDNA clone IMAGE:1741487.3 similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1 HOMOLOG (HUMAN); mRNA sequence.				
ACCESSION	AI193140				
NID	93744349				
VERSION	AI193140.1	GI:3744349			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.					
REFERENCE	1 (bases 1 to 475)				

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 JOURNAL Tumor Gene Index
 COMMENT Unpublished (1997)
 On Jan 14, 1998 this sequence version replaced gi:1797529.

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 911 Std Error: 0.00
 Seq primer: -40m3 fwd. ET from AmerSham
 High quality sequence stop: 329.
 Location/Qualifiers

FEATURES
 SOURCE
 1..475
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1741487"
 /clone_lib="Soares_fetal_Lung_NBHL19W"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: Lung; Vector: pT713D (Pharmacia) with a
 modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATCTGAAGTGGAGCGCGCATTTTCTTTTCTTTT-3']
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT713 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M. Fatima Bonaldo. This library was constructed
 from the same fetus as the fetal heart library, Soares
 fetal heart NBHL19W."
 BASE COUNT 176 a 101 c 73 g 125 t
 ORIGIN

Query Match 22.2%; Score 386; DB 43; Length 475;
 Best Local Similarity 100.0%; Pred. No. 3,5e-157;
 Matches 386: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1284 TTGTGAGTTTAATGACTATTATATATAATTTTGTGTTTCATATTGATGTGTGTC 1343
 |||||||
 DB 386 TTGTGAGTTTAATGACTATTATATATAATTTTGTGTTTCATATTGATGTGTGTC 327
 OY 1344 TGGGAGGACCTGTGGCCAGTTCTTACTGCTGTATGTCGTGTGTAGGACTGTGAAA 1403
 |||||||
 DB 326 TAGGAGGACCTGTGGCCAGTTCTTACTGCTGTATGTCGTGTGTAGGACTGTGAAA 267
 OY 1404 AGGGAAGTGAACATTCAGAGCGGTAGTGAATCAAGTAAAGCTAGAAATGCCACG 1463
 |||||||
 DB 266 AGGGAAGTGAACATTCAGAGCGGTAGTGAATCAAGTAAAGCTAGAAATGCCACG 207
 OY 1454 TGTATTATGATGATAATCTCCATCCCGTGGAAAGCTTTTCTGTTCTTAAGACGTG 1523
 |||||||
 DB 206 TGTATTATGATGATAATCTCCATCCCGTGGAAAGCTTTTCTGTTCTTAAGACGTG 147
 OY 1524 ATTTGCTGTAGAGATGACCTTATTAACCAAGGCCAAGGTGTATAAATGCTGGTT 1583
 |||||||
 DB 146 ATTTGCTGTAGAGATGACCTTATTAACCAAGGCCAAGGTGTATAAATGCTGGTT 87
 OY 1584 TTTCAGTTTTCAGAGTGGGTGATTTTCAGACCTACAGTGTACAGTCTTGTATTAGTT 1643
 |||||||
 DB 86 TTTCAGTTTTCAGAGTGGGTGATTTTCAGACCTACAGTGTACAGTCTTGTATTAGTT 27
 OY 1644 GTTATAAAGTACATGTTAACTTA 1669
 |||||||
 DB 26 GTTATAAAGTACATGTTAACTTA 1

RESULT 31
 A1434652/c

LOCUS A1434652 531 bp mRNA EST 30-MAR-1999
 DEFINITION ti34d02.x1 NCI-CGAP_Lym12 Homo sapiens cDNA clone IMAGE:213235 3'
 Similar to gb:D06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
 HOMOLOG (HUMAN); mRNA sequence.
 ACCESSION A1434652
 NID 94297677
 VERSION A1434652.1 GI:4297677
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 531)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On May 7, 1998 this sequence version replaced gi:312121.

Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Life Technologies catalog #: 11547-015
 DNA sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/dbp/image/image.html

Insert Length: 742 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 293.
 Location/Qualifiers

FEATURES
 SOURCE
 1..531
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="12p"
 /clone="IMAGE:2132355"
 /clone_lib="NCI-CGAP_Lym12"
 /tissue_type="lymphoma, follicular mixed small and large
 cell"
 /lab_host="DH10B"
 /note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:
 SalI; Site_2: NotI; Cloned unidirectionally. Primer:
 oligo dT. Average insert size 1.25 kb. Life Technologies
 catalog #: 11547-015"
 BASE COUNT 194 a 105 c 73 g 158 t 1 others
 ORIGIN

Query Match 21.6%; Score 375; DB 45; Length 531;
 Best Local Similarity 100.0%; Pred. No. 2e-152;
 Matches 375: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1316 TTTTGTGTTCAATTGATGTGTCTAGGACGCTGTGGCAAGTTCTTAGTGC 1375
 |||||||
 DB 375 TTTTGTGTTCAATTGATGTGTCTAGGACGCTGTGGCAAGTTCTTAGTGC 316
 OY 1376 TGTATGTCGTGGTGAAGCTAGAAAAGGAACTGAACATTCAGAGCGGTAGTGA 1435
 |||||||
 DB 315 TGTATGTCGTGGTGAAGCTAGAAAAGGAACTGAACATTCAGAGCGGTAGTGA 256
 OY 1436 TCACGTAAGCTAGAAATGATCCAGCTGTTATGCAATGATCTCCATTCCTG 1495
 |||||||
 DB 255 TCACGTAAGCTAGAAATGATCCAGCTGTTATGCAATGATCTCCATTCCTG 196
 OY 1496 GGAAGCTTTTCTGTTCTTAAGACGTGATTTTGTGTGAAGATGGCACTTAACCA 1555
 |||||||
 DB 195 GGAAGCTTTTCTGTTCTTAAGACGTGATTTTGTGTGAAGATGGCACTTAACCA 136
 OY 1556 AGCCCAAAAGTGTATGAAGATCTGTTTTCAGTTTCAGAGTGGGTGATTCAGCA 1615
 |||||||
 DB 135 AGCCCAAAAGTGTATGAAGATCTGTTTTCAGTTTCAGAGTGGGTGATTCAGCA 76

OY 1516 CCTACAGTGTACAGTCTTGTATTAGTTGTTAATAAAGTACATGTTAACTTAAAAA 1675
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 75 CCTACAGTGTACAGTCTTGTATTAGTTGTTAATAAAGTACATGTTAACTTAAAAA 16
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 1576 AAAAAAAAAAAAAA 1690
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 15 AAAAAAAAAAAAAA 1
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 32
 AA824607 538 bp mRNA EST 18-MAR-1998
 LOCUS OC83f08.s1 NCI_CGAP_GCB1 Homo sapiens cdna clone IMAGE:1356327
 DEFINITION similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
 HOMOLOG (HUMAN);, mRNA sequence.
 ACCESSION AA824607
 NID 92886629
 VERSION AA824607.1 GI:2896629
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 538)
 REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 TITLE Unpublished (1997)
 JOURNAL On Sep 29, 1997 this sequence version replaced gi:1520732.
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
 Ph.D., Gerald Marti, M.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 www-bio.llnl.gov/bdrrp/image/image.html
 Insert Length: 643 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 191.
 Location/Qualifiers
 1..538
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="1356327"
 /clone_lib="NCI_CGAP_GCB1"
 /tissue_type="germinal center B cell"
 /lab_host="DH10B"
 /note="vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from human tonsillar cells enriched for
 germinal center B cells by flow sorting (CD20+ IgD-),
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
 primed with a Not I - oligo(dT) primer
 15'-TGTACCAATCTGAAGTGGGCGCCGCTTTTCTTTTCTTTT-
 3'. Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 185 a 102 c 85 g 166 t
 ORIGIN
 Query Match 21.1%; Score 366; DB 39; Length 538;
 Best Local Similarity 100.0%; Pred. No. 1.5e-148;

Matches 366; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1325 TTTCATTTGATGTGTGTCTAGGACGACCTGGCCAGTTCATGTTGCTATGCT 1384
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 366 TTTCATTTGATGTGTGTCTAGGACGACCTGGCCAGTTCATGTTGCTATGCT 307
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 1385 CGTGTAGACCTGTAGAAAAGGAAGTCACTCCAGAGCGGTAGTGAATCAGCTAAA 1444
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 306 CGTGTAGACCTGTAGAAAAGGAAGTCACTCCAGAGCGGTAGTGAATCAGCTAAA 247
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 1445 GGTAGAAATGATCCCGAGCTGTTTATGATATATCTCTCATTCCTCCGTTGAAGTTT 1504
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 246 GGTAGAAATGATCCCGAGCTGTTTATGATATATCTCTCATTCCTCCGTTGAAGTTT 187
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 1505 TTCCTGTTCTTAAAGACGATTTTGGCTGAGAAAGTGCACCTTAATACCAAGCCCAAG 1564
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 186 TTCCTGTTCTTAAAGACGATTTTGGCTGAGAAAGTGCACCTTAATACCAAGCCCAAG 127
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 1565 TGGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTGTGATTTTCAGCACCCTACAGTG 1624
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 126 TGGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTGTGATTTTCAGCACCCTACAGTG 67
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 1625 TACAGTCTGTATTAGTTGTTTAAAGTACATCTTAACCTTAAAAA 1684
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 66 TACAGTCTGTATTAGTTGTTTAAAGTACATCTTAACTTAAAAA 7
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 1685 AAAAAA 1690
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 6 AAAAAA 1
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 33
 AA479357/c 537 bp mRNA EST 08-AUG-1997
 LOCUS 2v17f10.s1 Soares NhlMPu.S1 Homo sapiens cdna clone IMAGE:753931 3'
 DEFINITION similar to gb:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
 HOMOLOG (HUMAN);, mRNA sequence.
 ACCESSION AA479357
 NID 92207913
 VERSION AA479357.1 GI:2207913
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 537)
 REFERENCE Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
 Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theisling, B.,
 White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-Merck EST Project 1997
 TITLE Unpublished (1997)
 JOURNAL On Sep 12, 1996 this sequence version replaced gi:1393855.
 COMMENT
 Contact: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 372.
 Location/Qualifiers
 1..537
 /organism="Homo sapiens"
 /db_xref="GDB:5976803"
 /db_xref="taxon:9606"
 /map="19p12-p13.1"
 /clone_image="IMAGE:753931"
 /clone_lib="Soares_NhlMPu.S1"
 /tissue_type="Pooled human melanocyte, fetal heart, and

pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT73D-Bac
(pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NHM, pregnant uterus
NBHV, and fetal heart NBH19M) were mixed and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT 194 a 107 c 89 g 147 t
ORIGIN

Query Match 20.4%; Score 354; DB 34; Length 537;
Best Local Similarity 100.0%; Pred. No. 2.4e-143;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1318 TTTTGTTCATATGATGCTGCTAGCAGAGACTGTGCCAAGTTTATGTTGCTG 1377
|||||
Db 354 TTTTGTTCATATGATGCTGCTAGCAGAGACTGTGCCAAGTTTATGTTGCTG 295

QY 1378 TATGCTCGTAGAGACTGTAGAAAAGGAGACTGAACATTCAGAGCGGTGATGATC 1437
|||||
Db 294 TATGCTCGTAGAGACTGTAGAAAAGGAGACTGAACATTCAGAGCGGTGATGATC 235

QY 1438 ACCTAAGCTAGAAATGATCCCGAGCTGTTATGATAGATATCTCCATTCGCGTGG 1497
|||||
Db 234 ACCTAAGCTAGAAATGATCCCGAGCTGTTATGATAGATATCTCCATTCGCGTGG 175

QY 1498 AACGTTTTCTGTTCTTAAAGAGTATTTGCTGTAGAAAGAGGCACTTATACCAAG 1557
|||||
Db 174 AACGTTTTCTGTTCTTAAAGAGTATTTGCTGTAGAAAGAGGCACTTATACCAAG 115

QY 1558 CCCAAGTGTATAGAAATGCTGTTTTCAGTTTCAGAGAGTGGTGTATTCAGACAC 1617
|||||
Db 114 CCCAAGTGTATAGAAATGCTGTTTTCAGTTTCAGAGAGTGGTGTATTCAGACAC 55

QY 1618 TACAGTGTACAGTCTGTATTAAGTTTAAAGTATTAAGTATTAAGTATTAAGTATTAAG 1671
|||||
Db 54 TACAGTGTACAGTCTGTATTAAGTTTAAAGTATTAAGTATTAAGTATTAAGTATTAAG 1

RESULT 34
A1682902/c 405 bp mRNA EST 26-MAY-1999
LOCUS wc69g07.x1 NCI-CGAP_Pauli Homo sapiens cDNA clone IMAGE:233932 3'
DEFINITION similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
HOMOLOG (HUMAN); mRNA sequence.
ACCESSION A1682902
NID 94893084
VERSION A1682902.1 GI:4893084
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 405)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On May 18, 1998 this sequence version replaced gi:3137611.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html
Seq primer: -40UP from Glibco
High quality sequence stop: 339.
Location/Qualifiers
1.405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="X"
/clone="IMAGE:233932"
/clone_lib="NCI CGAP_Pauli"
/tissue_type="adrenocortihoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.72 kb. Life Technologies catalog #:
11548-013"

BASE COUNT 149 a 88 c 61 g 106 t 1 others
ORIGIN

Query Match 20.4%; Score 354; DB 50; Length 405;
Best Local Similarity 99.8%; Pred. No. 2.6e-143;
Matches 404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1266 TCTTGTGTTCTTAACTTTTGTGAAGTTAATGACTATTATATAATTTTGTGT 1325
|||||
Db 405 TCTTGTGTTCTTAACTTTTGTGAAGTTAATGACTATTATATAATTTTGTGT 346

QY 1326 TTCAATATGATGTTGTTCTAGAGAGACCGTGGCCAGTTCTTACTGTTATGCTC 1385
|||||
Db 345 TTCAATATGATGTTGTTCTAGAGAGACCGTGGCCAGTTCTTACTGTTATGCTC 286

QY 1386 GTGTAGAGACTGTAGAAAAGGAGAACTGAACATTCAGAGCGGTGATGATCAGTAAAG 1445
|||||
Db 285 GTGTAGAGACTGTAGAAAAGGAGAACTGAACATTCAGAGCGGTGATGATCAGTAAAG 226

QY 1446 CTGAAATATGCCCGAGCTGTTATGATAGATATCTCCATTCGCGTGGAGCTTTT 1505
|||||
Db 225 CTGAAATATGCCCGAGCTGTTATGATAGATATCTCCATTCGCGTGGAGCTTTT 166

QY 1506 TCCTGTTCTTAAAGCTGATTTTGTGTAGAGATGGCACTTTAAACCAAGCCCAAGT 1565
|||||
Db 165 TCCTGTTCTTAAAGCTGATTTTGTGTAGAGATGGCACTTTAAACCAAGCCCAAGT 106

QY 1566 GGTATAGAAATGCTGTTTTCAGTTTTCAGAGAGTGGTGTATTCAGACCTACAGTGT 1625
|||||
Db 105 GGTATAGAAATGCTGTTTTCAGTTTTCAGAGAGTGGTGTATTCAGACCTACAGTGT 46

QY 1626 ACACTGTTGATTAAGTTTGTATTAATAAAGTATTAAGTATTAAGTATTAAGTATTAAG 1670
|||||
Db 45 ACACTGTTGATTAAGTTTGTATTAATAAAGTATTAAGTATTAAGTATTAAGTATTAAG 1

RESULT 35
AA707668 786 bp mRNA EST 24-DEC-1997
LOCUS zj29f12.s1 Soares_fetal_liver spleen, INFUS.S1 Homo sapiens cDNA
DEFINITION clone IMAGE:451727 3' similar to gb:L06797 PROBABLE G
PROTEIN-COUPLED RECEPTOR LCR1 HOMOLOG (HUMAN); mRNA sequence.
ACCESSION AA707668
NID 92717586
VERSION AA707668.1 GI:2717586
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 786)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S.,
Kizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, K., Stepoe, M., Tan, F.,

TITLE
JOURNAL
COMMENT

Thising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-MCI human EST Project
Unpublished (1997)
On Nov 6, 1997 this sequence version replaced g1:315382.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 268.

FEATURES

source
1. 786
/organism="Homo sapiens"
/db_xref="GDB:138803"
/db_xref="taxon:9606"
/map="17:15q26.1"
/clone="IMAGE:451727"
/sex="male"
/sex="female"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pRT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
This is a subcloned version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - Oligo(dT) primer [5',
AACTGAGAGATTATTAAGACATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pRT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT
256 a 156 c 163 g 211 t

ORIGIN

Query Match
Best Local Similarity 99.6%; Score 347; DB 31; Length 786;
Matches 447; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1220 AAGACTGACCAATATTTGACAGTTTATTCCTGTTGGATTGTTGTTCTTT 1279
|||||
Db 449 AAGACTGACCAATATTTGACAGTTTATTCCTGTTGGATTGTTGTTCTTT 390
QY 1280 AGTTTGTGAGTTTAACTATTTATTAATAATTTTGTGTTTTCATATGATGTG 1339
|||||
Db 389 AGTTTGTGAGTTTAACTATTTATTAATAATTTTGTGTTTTCATATGATGTG 330
QY 1340 TGTCTAGGACGACCTGTGGCCAACTTTAGTTCTGTATGTCGCGTGAAGCTGA 1399
|||||
Db 329 TGTCTAGGACGACCTGTGGCCAACTTTAGTTCTGTATGTCGCGTGAAGCTGA 270
QY 1400 GAAAAGGAACTGACCATTTCCAGAGCGTGTAGTGAATCAGTAAAGCTAGAAATATCC 1459
|||||
Db 269 GAAAAGGAACTGACCATTTCCAGAGCGTGTAGTGAATCAGTAAAGCTAGAAATATCC 210
QY 1460 CAGCTGTTATGATGATATCTCTCCATTCCTGTAAGCTTTTCTGTTCTTAAGA 1519
|||||
Db 209 CAGCTGTTATGATGATATCTCTCCATTCCTGTAAGCTTTTCTGTTCTTAAGA 150
QY 1520 CGTGAATTTGCTGTAGAGAGTGGCACTTATTAACCAAGCCCAAGTGTATGAATGCT 1579
|||||
Db 149 CGTGAATTTGCTGTAGAGAGTGGCACTTATTAACCAAGCCCAAGTGTATGAATGCT 90
QY 1580 GGTGTTTTCAGTTTTCAGAGAGTGGGTTGATTTCAGACCTACAGTGTATGATTA 1639
|||||
Db 89 GGTGTTTTCAGTTTTCAGAGAGTGGGTTGATTTCAGACCTACAGTGTATGATTA 30

QY 1640 ACTGTGTTATATAAAGTACATGTTAACTT 1668
Db 29 ACTGTGTTATATAAAGTACATGTTAACTT 1

RESULT 36
AA284857 454 bp mRNA EST 08-AUG-1997
LOCUS z122e05.s1 Soares ovary tumor Nshot Homo sapiens cDNA clone
DEFINITION IMAGE:713888 3' similar to gb:L06797 PROBABLE G PROTEIN-COUPLED
RECEPTOR LCRI HOMOLOG (HUMAN);, mRNA sequence.

ACCESSION
AA284857
NID 91927416
VERSION AA284857.1 GI:1927416
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chapell,B.,
Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
Roifling,T., Scheinberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.
and Marra,M.

TITLE
JOURNAL
MEDLINE
COMMENT

Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
On May 9, 1995 this sequence version replaced g1:802981.

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 561 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 294.

FEATURES

source
Location/Qualifiers

1. 454
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:713888"
/clone_lib="Soares ovary tumor NBHOT"
/sex="female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: ovary; Vector: pRT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - Oligo(dT) primer [5',
TGTTCACCACTGAGAGGAGCGCGCGGCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pRT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."

BASE COUNT
159 a 95 c 66 g 134 t

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 342; DB 31; Length 454;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1346 GGCAGACCTGTGGCCAAAGTTCTTATGTTGCTGTATGTCGTTGAGTGTAGAAAG 1405
|||||
Db 348 GGCAGACCTGTGGCCAAAGTTCTTATGTTGCTGTATGTCGTTGAGTGTAGAAAG 289

Oy	1406	GGAAGTGAACATTCACAGACCCTGTAATGCATCACTAAAGCTGAATGATCCCGACTG	1465
Db	288	GGAAGTGAACATTCACAGACCCTGTAATGCATCACTAAAGCTGAATGATCCCGACTG	229
Oy	1466	TTATGATAGATAAATCTCTCCATTCCGTGGAGAGCTTTTCTCCTGTTCTTAGACGTGAT	1525
Db	228	TTATGATAGATAAATCTCTCCATTCCGTGGAGAGCTTTTCTCCTGTTCTTAGACGTGAT	169
Oy	1526	TTTGCTGTAGAAGATGGCCACTTAATACCAAAGCCAAAGGTATGAATGCTGCTGTTT	1585
Db	168	TTTGCTGTAGAAGATGGCCACTTAATACCAAAGCCAAAGGTATGAATGCTGCTGTTT	109
Oy	1586	TCACGTTTTACAGAGTGGGTGATTGATTTTCAGCACCCATACAGTGTACAGTCTTGATATTAAAGTTG	1645
Db	108	TCACGTTTTACAGAGTGGGTGATTGATTTTCAGCACCCATACAGTGTACAGTCTTGATATTAAAGTTG	49
Oy	1646	TAAATAAGTACATGTTAACCTTAACAAAAAAAAAAAAAAA 1687	
Db	48	TAAATAAGTACATGTTAACCTTAACAAAAAAAAAAAAAAA 7	
RESULT	37		
LOCUS	AA284569		
DEFINITION	AA284569 494 bp mRNA EST 08-AUG-1997		
ACCESSION	Z122605.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone IMAGE:713888 5' similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI HOMOLOG (HUMAN);, mRNA sequence.		
NID	AA284569		
VERSION	g1927533		
KEYWORDS	AA284569.1 GI:1927533		
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
AUTHORS	1 (bases 1 to 494) Hillier,L., Lennon,G., Becker,M., Bonaldo M.F., Chapelli,B., Chissee,S., Dietrich,N., Dubouque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mandis,E., Moore,B., Morris,B., Parsons,J., Piarange,C., Rifkin,L., Rohlfing,T., Scheilenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevaskis,F., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.		
TITLE	Generation and analysis of 280,000 human expressed sequence tags		
JOURNAL	Genome Res. 6 (9), 807-828 (1996)		
MEDLINE	97044478		
COMMENT	On May 9, 1995 this sequence version replaced gi:804201.		
FEATURES	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel.: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 561 Std Error: 0.00 Seq primer: -28m3 rev2 ET from Amersham High quality sequence stop: 383. Location/Qualifiers		
SOURCE	1..494 /organism="Homo sapiens" /db_xref="taxon:9606" /map="16" /clone="IMAGE:713888" /clone_lib="Soares ovary tumor NbHOT" /sex="female" /tissue_type="ovarian tumor" /lab_host="DH10B (ampicillin resistant)" /note="Organ: ovary; Vector: pTR10 (pharmacía) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' GTTACCATCTGAGTGGAGCCGCCGGGCTTTTCTTTTCTTTT 3'] ,		

Query Match	18.9%	Score 328	DB 31	Length 494
Best Local Similarity	99.7%	Pred. No. 4.4e-132		
Matches 378	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY 1189	TTTTTTTAAGTACACATTTTTCAGATATAAAGACGTGCACATATTTGACAGTTTAT	1248		
Db 1	TTTTTTTAAGTACACATTTTTCAGATATAAAGACGTGCACATATTTGACAGTTTAT	60		
QY 1249	TGCTTTGTTGGATTTTGTCTTGTGTTCTTTAGTTTGTGAAGTTAATGACTTATTT	1308		
Db 61	TGCTTTGTTGGATTTTGTCTTGTGTTCTTTAGTTTGTGAAGTTAATGACTTATTT	120		
QY 1309	ATAATAATTTTGTGTTTCATATGATGTGCTGTACGCAGACCTGTGGCCAGTTCT	1368		
Db 121	ATAATAATTTTGTGTTTCATATGATGTGCTGTACGCAGACCTGTGGCCAGTTCT	180		
QY 1369	TAGTTGCTGTATGCTCGTGAGACACTGTAGAAAAGGGAAGTGAACAATTCAGACGCTG	1428		
Db 181	TAGTTGCTGTATGCTCGTGAGACACTGTAGAAAAGGGAAGTGAACAATTCAGACGCTG	240		
QY 1429	TAGTGAATCAGCTAAAGCTAGAAATGATGCCAGCTGTTTATGCAATGATATCTCTCA	1488		
Db 241	TAGTGAATCAGCTAAAGCTAGAAATGATGCCAGCTGTTTATGCAATGATATCTCTCA	300		
QY 1489	TTCCCGTGGAAAGCTTTTCCCTGTTCTTTAGACAGCTATTTGCTGTAGAAGATGGACCTTA	1548		
Db 301	TTCCCGTGGAAAGCTTTTCCCTGTTCTTTAGACAGCTATTTGCTGTAGAAGATGGACCTTA	360		
QY 1549	TAACCAAAGCCCAAGTGG 1567			
Db 361	TAACCAAAGCCCAAGTGG 379			
RESULT 38				
AA501992/c				
LOCUS	AA501992	308 bp	mRNA	EST
DEFINITION	ncs1h08.s1 NCI-CGAP_C03 Homo sapiens cDNA clone IMAGE:900927 3'			
	similar to gp:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1			
	HOMOLOG (HUMAN), mRNA sequence.			
ACCESSION	AA501992			
NID	92236959			
VERSION	AA501992.1	GI:2236959		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
	Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 308)			
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .			
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),			
	Tumor Gene Index			
JOURNAL	Unpublished (1997)			
COMMENT	On Sep 12, 1996 this sequence version replaced gi:1398118.			

double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Donald.

Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E Consortium/LLNL at:

www.bio.lnl.gov/bdrp/image/image.html

Insert Length: 418 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham.

Location/Qualifiers
1. 308

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:900927"
/clone_1lb="NCI_CGAP_Co3"
/sex="pooled"
/tissue.type="colon"
/lab.host="DH10B"

/note="vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization."

BASE COUNT 100 a 69 c 52 g 87 t
ORIGIN

Query Match 17.7% Score 308; DB 34; Length 308;
Best Local Similarity 100.0%; Pred. No. 2,3e-123;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 CAAGTCTTAGTGTGCTGATGCTGCTGAGAGCTGTAGAAAAGGGAAGTCC 1420
DB 308 CAAGTCTTAGTGTGCTGATGCTGCTGAGAGCTGTAGAAAAGGGAAGTCC 249

QY 1421 AAGAGCTGTAGTATCAGTATGAGTAAAGATGCCAGCTGTATGATGATGATA 1480
DB 248 AAGAGCTGTAGTATGAGTAAAGATGCCAGCTGTATGATGATGATA 189

QY 1481 TCTCTCATTCCTGCTGAGAGCTTTTCTCTTCTTAAGAGTATTTGCTGTAGAGAT 1540
DB 188 TCTCTCATTCCTGCTGAGAGCTTTTCTCTTCTTAAGAGTATTTGCTGTAGAGAT 129

QY 1541 GGCACCTTAACCAAGCCCAAGTGTATAGAAATGCTGTTTCAGTTTCAGAGAT 1600
DB 128 GGCACCTTAACCAAGCCCAAGTGTATAGAAATGCTGTTTCAGTTTCAGAGAT 69

QY 1601 GGGTATTTACACCTACAGTGTACAGTCTTGTATTAAGTTGTTAATAAAGTACATG 1660
DB 68 GGGTATTTACACCTACAGTGTACAGTCTTGTATTAAGTTGTTAATAAAGTACATG 9

QY 1661 TTAAGCT 1668
DB 8 TTAAGCT 1

RESULT 39 EST 21-APR-1999
LOCUS A1610565/c 361 bp mRNA
DEFINITION tp41808.x1 NCI_CGAP ut4 Homo sapiens cDNA clone IMAGE:2190351 3'
HOMOLOG (HUMAN): mRNA sequence.

ACCESSION A1610565
NID 94619732
VERSION A1610565.1 GI:4619732
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 361)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)

COMMENT

On Mar 10, 1998 this sequence version replaced gi:2949515.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.lnl.gov/bdrp/image/image.html

FEATURES

Source

Seq primer: -40UP from Gldco
High quality sequence stop: 219.
Location/Qualifiers

1. 361
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2190351"
/clone_1lb="NCI_CGAP_Ut4"
/tissue.type="serous papillary carcinoma, high grade, 2
pooled tumors"
/lab.host="DH10B"

BASE COUNT 120 a 82 c 58 g 101 t
ORIGIN

Query Match 17.2% Score 299; DB 48; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.7e-119;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1373 TGTGTATGTCTGCTGAGTGTAGAAAAGGGAAGTCCAGAGGTGTACT 1432
DB 299 TGTGTATGTCTGCTGAGTGTAGAAAAGGGAAGTCCAGAGGTGTACT 240

QY 1433 GATTCAGTAAAGCTGAATATGCCAGCTGTTATCATAGATATCTCCATTCC 1492
DB 239 GATTCAGTAAAGCTGAATATGCCAGCTGTTATCATAGATATCTCCATTCC 180

QY 1493 CGTGAACGTTTTTCTGTTTAAAGCGTATTTGCTGTAGAAATGCGACTTTAAC 1552
DB 179 CGTGAACGTTTTTCTGTTTAAAGCGTATTTGCTGTAGAAATGCGACTTTAAC 120

QY 1553 CAAGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTCAGAGTGGTGTATTCA 1612
DB 119 CAAGCCCAAGTGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGTGTATTCA 60

QY 1613 GCACCTACAGTGTACAGTCTTGTATTAGTTGTTAATAAAGTATTAACTTAA 1671
DB 59 GCACCTACAGTGTACAGTCTTGTATTAGTTGTTAATAAAGTATTAACTTAA 1

RESULT 40 EST 16-FEB-1995
LOCUS T62636/c 407 bp mRNA
DEFINITION yc03f11.s1 Strathgane lung (#937210) Homo sapiens cDNA clone
IMAGE:79629 3' similar to gb:L06797 PROBABLE G PROTEIN-COUPLED
RECEPTOR LCRI HOMOLOG (HUMAN): mRNA sequence.

ACCESSION T62636
NID 9666293
VERSION T62636.1 GI:666293
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 407)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B., Chisnoe, S., Dietrich, N., Dubuque, T., Favellio, A., Gish, M., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Madris, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, I., Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J., Trevasakis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R., and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags

JOURNAL Genome Res. 6 (9), 807-828 (1996)

MEDLINE 97044478

COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 333
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -21ml3
High quality sequence stop: 333.
Location/Qualifiers
1. 407
/organism="Homo sapiens"
/db_xref="GDB:483246"
/db_xref="taxon:9606"
/clone="IMAGE:79629"
/clone_lib="Stratagene lung (#937210)"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site_1:
ECORI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. normal lung. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'."

BASE COUNT 141 a 88 c 65 g 112 t 1 others

ORIGIN

Query Match 17.2%; Score 298; DB 21; Length 407;
Best Local Similarity 100.0%; Pred. No. 4.5e-119;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1367 CTTAGTTCCTGATGTCCTGCTGAGACTGTAGAAAGGGAAGTGAACATTCAGAGCG 1426
|||||
Db 306 CTTAGTTCCTGATGTCCTGCTGAGACTGTAGAAAGGGAAGTGAACATTCAGAGCG 247

QY 1427 TGTAGTGAATCAGCTAAGCTGAAGTATGATCCCGAGCTGTTATGCATGATATCTCTC 1486
|||||
Db 246 TGTAGTGAATCAGCTAAGCTGAAGTATGATCCCGAGCTGTTATGCATGATATCTCTC 187

QY 1487 CATTCGCGTGAAGCTTTTCTCTGTTCTTAAGAGCTGATTTGCTGTAGAGATGGCACT 1546
|||||
Db 186 CATTCGCGTGAAGCTTTTCTCTGTTCTTAAGAGCTGATTTGCTGTAGAGATGGCACT 127

QY 1547 TATAACCAAGCCCAAGTGTGATGAAGTGTGTTTTCGTTTTCAGTGAGTGCGTTG 1606
|||||
Db 126 TATAACCAAGCCCAAGTGTGATGAAGTGTGTTTTCGTTTTCAGTGAGTGCGTTG 67

QY 1607 ATTTCAGACCTACAGTGTAGAGTCTGTATTAAGTTGTTAATAAAGTACATGTTAA 1664
|||||
Db 66 ATTTCAGACCTACAGTGTAGAGTCTGTATTAAGTTGTTAATAAAGTACATGTTAA 9

RESULT 41
AA621854 342 bp mRNA EST 21-OCT-1997
LOCUS nq19901.s1 NCI CGAP, Thy1 Homo sapiens cDNA clone IMAGE:1144368
DEFINITION Similar to DB:106797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCR1
HOMOLOG (HUMAN):, mRNA sequence.

ACCESSION AA621854
NID 92525730
VERSION AA621854.1 GI:2525730
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 342)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1407054.

CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Kriman, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/1image/image.html

FEATURES
SOURCE
1. 342
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1144368"
/clone_lib="NCI-CGAP, Thy1"
/tissue_type="thyroid"
/lab_host="DH10B"
/note="Vector: pAMP10; mRNA made from invasive thyroid
tumor, cDNA made by oligo-dt priming. Non-directionally
cloned. Size-selected on agarose gel, average insert size
600 bp. Reference: Krizman et al. (1996) Cancer Research
56:5380-5383."

BASE COUNT 78 a 105 c 73 g 86 t

ORIGIN

Query Match 17.0%; Score 296; DB 36; Length 342;
Best Local Similarity 100.0%; Pred. No. 3.4e-118;
Matches 296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 778 AAGGCGCACCGAAGCGCAAGGCCCTCAAGACCAAGTATCTATCTGCTTTCTTC 837
|||||
Db 12 AAGGCGCACCGAAGCGCAAGGCCCTCAAGACCAAGTATCTATCTGCTTTCTTC 71

QY 838 GCGCTTTGGCTGCTTACTACATTGGAGTACAGATCGATCCTTCATCCTCGGAATC 897
|||||
Db 72 GCGCTTTGGCTGCTTACTACATTGGAGTACAGATCGATCCTTCATCCTCGGAATC 131

QY 898 ATGANGCAAGGGGTGTGAGTTTGAGAACACTGTGCAAGTGGATTTTCATCACCGAGGCC 957
|||||
Db 132 ATGANGCAAGGGGTGTGAGTTTGAGAACACTGTGCAAGTGGATTTTCATCACCGAGGCC 191

QY 958 CTAGCTTCTTCACATGTGCTGTGAACCCATCCCTATGCTTCTTGAGGCCAATTT 1017
|||||
Db 192 CTAGCTTCTTCACATGTGCTGTGAACCCATCCCTATGCTTCTTGAGGCCAATTT 251

QY 1018 AAAACCTTGCCACGACGACCTACCTCTGTGAGCAGAGGGTCCAGCCTCAAGAT 1073
|||||
Db 252 AAAACCTTGCCACGACGACCTACCTCTGTGAGCAGAGGGTCCAGCCTCAAGAT 307

RESULT 42

Thu Aug 12 09:17:07 1999

us-09-104-063-3.1.txt

Page 34

AA386001
LOCUS AA386001 438 bp mRNA EST 21-Apr-1997
DEFINITION E5199759 Pancreas tumor I Homo sapiens cDNA 5' end similar to
guanine nucleotide-binding protein HM89, mRNA sequence.
ACCESSION AA386001
NID 92038338
VERSION AA386001.1 GI:2038338
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 438)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Wal,C.,
Wille,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Wal,C.,
Clayton,R.A., Cline,T.R., Colton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glocke,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaro,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligriano,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudex,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,G.A.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M., and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
JOURNAL 96026280
MEDLINE On Dec 30, 1996 this sequence version replaced gi:1530938.
COMMENT Other-ESTs: E5199758 TRC166975
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@igrr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.
Location/Qualifiers
FEATURES
Source 1. 438
/organism="Homo sapiens"
/db_xref="ATCC (inhost):190162"
/db_xref="taxon:9606"
/clone_lib="Pancreas tumor I"
/dev_stage="adult"
/note="organ: pancreas; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
BASE COUNT 104 a 104 c 83 g 145 t 2 others
ORIGIN
Query Match 17.0%; Score 295; DB 33; Length 438;
Best Local Similarity 100.0%; Pred. No. 8,7e-118;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 121 TTCCAGTGTGTGAGACCCCATCTCTATCTTCTTGAGCCAAATTAACCTCT 180
Oy 1027 GCCCAGCAGCAGCTACCTCTGTGAGCAGAGGGTTCAGCCCTCAAGATCTCTCCAAAGGA 1086
Db 181 GCCCAGCAGCAGCTACCTCTGTGAGCAGAGGGTTCAGCCCTCAAGATCTCTCCAAAGGA 240
Oy 1087 AAGCAGGTGACATTCATCTGTTCACCTGAGTGTGAGTTCAGTTTCACT 1141
Db 241 AAGCAGGTGACATTCATCTGTTCACCTGAGTGTGAGTTCAGTTTCACT 295
RESULT 43
AA354393 456 bp mRNA EST 21-Apr-1997
LOCUS AA354393
DEFINITION E5162864 Jurkat T-cells V Homo sapiens cDNA 5' end similar to
guanine nucleotide-binding protein HM89, mRNA sequence.
ACCESSION AA354393
NID 92006897
VERSION AA354393.1 GI:2006897
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 456)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Wal,C.,
Clayton,R.A., Cline,T.R., Colton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glocke,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaro,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligriano,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudex,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,G.A.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M., and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
JOURNAL 96026280
MEDLINE On May 9, 1995 this sequence version replaced gi:802440.
COMMENT Other-ESTs: TRC166975
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@igrr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.
Location/Qualifiers
FEATURES
Source 1. 456
/organism="Homo sapiens"
/db_xref="ATCC (inhost):154827"
/db_xref="taxon:9606"
/map="12P"
/clone_lib="Jurkat T-cells V"
/cell_type="T-lymphocyte"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 116 a 121 c 105 g 111 t 3 others
ORIGIN

Query Match 16.9%; Score 293; DB 32; Length 456;
Best Local Similarity 100.0%; Pred. No. 6,3e-117;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

59 TACCACCCGATGGAGACACGCGTTACCTGAGGGGATCATATATACACTTCAG 118
|||
33 TACCACCCGATGGAGACACGCGTTACCTGAGGGGATCATATATACACTTCAG 92
|||
119 ATATACACACGAGAAATGGCTCAGGGAGCTATGACTCATGAGAAACCTGTTCC 178
|||
93 ATACTACACCGAGAAATGGCTCAGGGAGCTATGACTCATGAGAAACCTGTTCC 152
|||
179 GTGAAGAAATGCTATTTCAATTAATATCTCTGCGCCACCATCTACTCATCTCT 238
|||
153 GTGAAGAAATGCTATTTCAATTAATATCTCTGCGCCACCATCTACTCATCTCT 212
|||
239 TAACGGCATTTGGGCAATGATGTTGTCATCTCTGTCATGGTTTACAGAGAACTGA 298
|||
213 TAACGGCATTTGGGCAATGATGTTGTCATCTCTGTCATGGTTTACAGAGAACTGA 272
|||
299 GAAGCATGACGCAAGTACAGGCTGACCTGTGAGTGGCCGACCTCTCTTT 351
|||
273 GAAGCATGACGCAAGTACAGGCTGACCTGTGAGTGGCCGACCTCTCTTT 325
|||

RESULT 44
AA747545/c 555 bp mRNA EST 17-MAR-1999
LOCUS
DEFINITION
nxB5d08.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1269039 3'
similar to gb:U06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.

ACCESSION
NID
AA747545 92787503
VERSION
AA747545.1 GI:2787503
KEYWORDS
EST.
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 555)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 18, 1995 this sequence version replaced gi:811242.

JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

FEATURES
source
1. 555
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="20"
/clone="IMAGE:1269039"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pT713D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from human consiliar cells enriched for
germinal center B cells by flow sorting (CD20-, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTTCACATCGAGTGGAGCGGCCGCCCATTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 210 a 108 c 84 g 153 t

Query Match 16.8%; Score 292; DB 38; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.6e-116;
Matches 412; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1266 TCTTGTGTTCTTTAGTTTGTGGAAGTTAATTGACTTATTAATAATTTTGT 1325
|||
413 TCTTGTGTTCTTTAGTTTGTGGAAGTTAATTGACTTATTAATAATTTTGT 354
|||
1326 TTCAATATTGATGTGTGCTGAGGACGCTGCGCAAGTTCTAGTGTATGCTC 1385
|||
353 TTCAATATTGATGTGTGCTGAGGACGCTGCGCAAGTTCTAGTGTATGCTC 294
|||
1386 GTGTGAGGACTGTAGAAAAGGAACTGAACATTCAGAGCGTGTAGTAATCACTGAAG 1445
|||
293 GTGTGAGGACTGTAGAAAAGGAACTGAACATTCAGAGCGTGTAGTAATCACTGAAG 224
|||
1446 CTAGAAATGATCCCGACGCTGTTATGATGATTAATCTCTCCA-TTCCGTGGAAGCTT 1504
|||
233 CTAGAAATGATCCCGACGCTGTTATGATGATTAATCTCTCCA-TTCCGTGGAAGCTT 174
|||
1505 TTCTGTTCTTAAAGAGTGTATTTGCTGTGAGAGATGAGCATATTAACCAACCCCAAG 1564
|||
173 TTCTGTTCTTAAAGAGTGTATTTGCTGTGAGAGATGAGCATATTAACCAACCCCAAG 114
|||
1565 TGGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGGTTGATTTTCAGCCTACAGTG 1624
|||
113 TGGTATAGAAATGCTGTTTTCAGTTTTCAGAGTGGGTTGATTTTCAGCCTACAGTG 54
|||

1625 TACAGCTTGATTAAGTTGTTAATAAGTACATGTTAACTTAAATAAAAAA 1677
|||
53 TACAGCTTGATTAAGTTGTTAATAAGTACATGTTAACTTAAATAAAAAA 1

RESULT 45
A1583169/c 524 bp mRNA EST 06-APR-1999
LOCUS
DEFINITION
tr98d12.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2227127 3'
similar to gb:U06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.

ACCESSION
NID
A1583169 94569066
VERSION
A1583169.1 GI:4569066
KEYWORDS
EST.
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 524)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 18, 1998 this sequence version replaced gi:3136741.

JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov


```

RESULT 47
LOCUS A1582682/c 277 bp mRNA EST 06-APR-1999
DEFINITION tm16d04.x1 NCI-CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2167783 3'
similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION A1582682
NID 94568579
VERSION A1582682.1 GI:4568579
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 277)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute / National Institute of Neurological
TITLE Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BI GAP), Tumor Gene Index
JOURNAL Unpublished (1998)
COMMENT On May 7, 1998 this sequence version replaced gi:3121491.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/dbtrp/image/image.html

Seg primer: -40UP from Gibco.
Location/Qualifiers
1. .277
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2167783"
/clone_lib="NCI-CGAP_Brn25"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site1: Not I; Site2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCCATGATGTTTGTGTGTGTGTGTGT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 85 a 58 c 46 g 88 t
ORIGIN

Query Match 15.9%; Score 277; DB 48; Length 277;
Best Local Similarity 100.0%; Pred. No. 6.1e-110;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1400 GAAAGGGAACGACATTCAGAGCGGTGATGATCAGCTAAGCTAGAAATGATCCC 1459
|||||
DB 277 GAAAGGGAACGACATTCAGAGCGGTGATGATCAGCTAAGCTAGAAATGATCCC 218

QY 1460 CAGCGTTATGCAATGATATCTCTCCATTCCTCGAAGCTTTTCTGTTTAA 1519
|||||
DB 217 CAGCGTTATGCAATGATATCTCTCCATTCCTCGAAGCTTTTCTGTTTAA 158

QY 1520 CGTGATTTGCTGTAAGATGCGCACTATATACCAAGCCCAAGTGTATGAATGCT 1579
|||||
DB 157 CGTGATTTGCTGTAAGATGCGCACTATATACCAAGCCCAAGTGTATGAATGCT 98

```

```

QY 1580 GGTTTTCAGCTTTTCAGAGAGTGGGTGATTTCAGCACCCTACAGTGTCTGTATTA 1639
|||||
DB 97 GGTTTTCAGCTTTTCAGAGAGTGGGTGATTTCAGCACCCTACAGTGTCTGTATTA 38

QY 1640 AGTTGTTAATAAAGTACATGTTAACTTAATAAAAAA 1676
|||||
DB 37 AGTTGTTAATAAAGTACATGTTAACTTAATAAAAAA 1

RESULT 48
LOCUS AA934643/c 427 bp mRNA EST 07-JUL-1998
DEFINITION o071c08.s1 NCI-CGAP_GC4 Homo sapiens cDNA clone IMAGE:1571630 3'
similar to gb:L06797 PROBABLE G PROTEIN-COUPLED RECEPTOR LCRI
HOMOLOG (HUMAN);, mRNA sequence.
ACCESSION AA934643
NID 93091855
VERSION AA934643.1 GI:3091855
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 427)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 14, 1998 this sequence version replaced gi:1797648.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/dbtrp/image/image.html

Insert Length: 989 Std Error: 0.00
Seg primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 315.
Location/Qualifiers
1. .427
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1571630"
/clone_lib="NCI-CGAP_GC4"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

BASE COUNT 146 a 87 c 67 g 125 t
ORIGIN

Query Match 15.8%; Score 275; DB 40; Length 427;
Best Local Similarity 100.0%; Pred. No. 4e-109;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1410 CTGAACATTCAGAGCGGTGATGATCAGCTAAGCTAGAAATGATCCAGCTGTTA 1469
|||||
DB 276 CTGAACATTCAGAGCGGTGATGATCAGCTAAGCTAGAAATGATCCAGCTGTTA 217

```

```

FEATURES
SOURCE
    location/Qualifiers
    1..347
    /organism="Homo sapiens"
    /db_xref="GDB:565965"
    /db_xref="taxon:9606"
    /clone="IMAGE:153626"
    /clone_lib="Soares breast 2NBH8t"
    /sex="female"
    /dev_stage="adult"
    /lab_host="DH10B (ampicillin resistant)"
    /note="Organ: breast; Vector: pT73D (Pharmacia) with a
    modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
    strand cDNA was primed with a Not I - oligo(dT) primer [5'
    TGTTACCATCTGAAGTGGAGCGCGCCGCTTTTCTTTTCTTTT 3'],
    double-stranded cDNA was ligated to Eco RI adaptors
    (Pharmacia), digested with Not I and cloned into the Not I
    and Eco RI sites of a modified pT73 vector (Pharmacia).
    Library went through one round of normalization to a Cot
    230. Library constructed by Bento Soares and M.Fátima
    Bonaldo."

```

BASE COUNT	109 a	81 c	62 g	95 t
ORIGIN				
Query Match	15.8%	Score 274;	DB 22;	Length 347;
Best Local Similarity	100.0%;	Pred. No. 1..le-108;		
Matches 274;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY 1383	CTCGGTGAGCACTCTAGAAAAAGGCACTGCAACATCCAGAGCCGTGTGATGATCACGTA	1442		
Db 288	CTCGGTGAGCACTCTAGAAAAAGGCACTGCAACATCCAGAGCCGTGTGATGATCACGTA	229		
QY 1443	AAGCTAGAAATGATCCCGACGCTTTATGCAATGATATCTCCATTCGCCGTGAAGCT	1502		
Db 228	AAGCTAGAAATGATCCCGACGCTTTATGCAATGATATCTCCATTCGCCGTGAAGCT	169		
QY 1503	TTTTCCTGTTCTTAAAGCGTATTTTCTGTGAGAGATGCGACTTATAACCAAAAGCCCAA	1562		
Db 168	TTTTCCTGTTCTTAAAGCGTATTTTCTGTGAGAGATGCGACTTATAACCAAAAGCCCAA	109		
QY 1563	AGTGTATGAAATGCTGGTTTTCAGTTTCAGAGATGGGTTGATTTACGACCTACAG	1622		
Db 108	AGTGTATGAAATGCTGGTTTTCAGTTTTCAGAGATGGGTTGATTTACGACCTACAG	49		
QY 1623	TGTACAGTCTGTATTAAGTTGTTAATAAAGTA	1656		
Db 48	TGTACAGTCTGTATTAAGTTGTTAATAAAGTA	15		

```

RESULT      50
A1581657/c

LOCUS
DEFINITION
ACCESSION
NID
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

A1581657.1 274 bp mRNA EST 06-APR-1999
AS038054 x1 Barstead colon HPLRB7 Homo sapiens cDNA clone
IMAGE:2174071.3' similar to gb:106797 PROBABLE G PROTEIN-COUPLED
RECEPTOR LCRI HOMOLOG (HUMAN);, mRNA sequence.
A1581657
A1581654 94567554
A1581657.1 GI:4567554
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 274)
Hiller,L., Allen,M., Bowles,L., Dubnque,T., Geisels,G., Jost,S.,
Krizman,D., Kucuba,T., Lacey,M., Le,N., Lennon,G., Marra,M.,
Mertins,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST project
Unpublished (1997)
On Mar 20, 1998 this sequence version replaced gi:2980476.

Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel.: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from GIBCO.

Location/Qualifiers
1. 274
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:2174071"
/clone_id="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/notes="organ: colon; Vector: pTZ19-Pac (Pharmacia) with a
modified polylinker; Site:1: EcoRI; Site:2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5

```

TGTTACGAATCTGAAGTGGAGCGGCCCTTTTTTTTTTTTTTTTTTTT
 3'] : double-stranded cDNA was ligated to Eco RI adaptors
 [5' AATTCACCTAGTAAT 3' and 5' ATTACTAGTg 3'], digested
 with Not I and cloned into the Not I and Eco RI sites of
 the modified pT713 vector. Library constructed by Bob
 Barstead."

BASE COUNT 85 a 58 c 46 g 85 t
 ORIGIN

Query Match 15.8%: Score 274; DB 48; Length 274;
 Best Local Similarity 100.0%; Pred. No. 1.2e-108;
 Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1400 GAAAGGGAGACTGAACATTCAGAGCGGTAGTGAATCAGTAAGCTAAGTAAGATGCC 1459
 DB 274 GAAAGGGAGACTGAACATTCAGAGCGGTAGTGAATCAGTAAGCTAAGTAAGATGCC 215
 OY 1460 CAGCTGTTATGCAATAGATATCTCTCATTCCTCCGTGAACGTTTCTCTTCTTAGA 1519
 DB 214 CAGCTGTTATGCAATAGATATCTCTCATTCCTCCGTGAACGTTTCTCTTCTTAGA 155
 OY 1520 CGTGATTTTGGTGTAGAGATGGCACTTATTAACCAAGCCCAAGTGTATAGAAATGCT 1579
 DB 154 CGTGATTTTGGTGTAGAGATGGCACTTATTAACCAAGCCCAAGTGTATAGAAATGCT 95
 OY 1580 GGTITTTTCAGTTTTCAGAGTGGTGTGATTTTCAGCACCTACAGTGTACAGTCTTGTATTA 1639
 DB 94 GGTITTTTCAGTTTTCAGAGTGGTGTGATTTTCAGCACCTACAGTGTACAGTCTTGTATTA 35
 OY 1640 AGTTGTTAATAAAGTACATGTTAACTTAAAA 1673
 DB 34 AGTTGTTAATAAAGTACATGTTAACTTAAAA 1

Search completed: August 8, 1999, 02:51:16
 Job time: 2026 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 7, 1999, 21:24:15 ; Search time 13.25 Seconds

(without alignments)
629.246 Million cell updates/sec

Title: US-09-104-063-4
Perfect score: 352

Sequence: 1 MEGISITSDNYTEEMSGSD.....KRGHSSVSTESSSFHSS 352

Scoring table: OLIGO

Searched: 188963 seqs, 23686106 residues

Database: A_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	352	100.0	352	1 R68812	Human monocyte PF4
2	352	100.0	352	1 R80757	Chemokine superfamily
3	347	98.6	356	1 W97362	G-protein coupled
4	337	95.7	337	1 W48734	Human RM3 seven tr
5	273	77.6	352	1 R27792	New platelet factor
6	236	67.0	337	1 R53753	Seven transmembran
7	83	23.6	92	1 W97363	A partial CXCR4 p
8	60	17.0	359	1 W64778	A murine CXCR4 p
9	25	7.1	25	1 W34001	Fusin receptor ant
10	22	6.2	22	1 W34003	Fusin receptor ant
11	22	6.2	22	1 W34005	Fusin receptor ant
12	21	6.0	21	1 W34004	Fusin receptor ant
13	21	6.0	21	1 W34006	Fusin receptor ant
14	21	6.0	21	1 W34007	Fusin receptor ant
15	12	3.4	20	1 W34002	Fusin receptor ant
16	10	2.8	350	1 R27791	Interleukin-8 rece
17	10	2.8	355	1 R28272	Interleukin-8 rece
18	10	2.8	360	1 R28273	Sequence in a high
19	10	2.8	355	1 R33420	Human IL-8 recepto
20	10	2.8	18	1 R37903	Interleukin-8 rece
21	10	2.8	350	1 R68811	Interleukin-8 rece
22	10	2.8	367	1 R67671	Mouse opiod recept
23	10	2.8	367	1 R71968	Rat opiod recepto
24	10	2.8	367	1 R74298	Mouse kappa-3 opio
25	10	2.8	367	1 R76638	Rat opiorph recept
26	10	2.8	1060	1 R70123	IL8-R type 1-GBP 1
27	10	2.8	1064	1 R70124	IL8-R type 2-GBP 1
28	10	2.8	350	1 R80756	Interleukin 8 rece
29	10	2.8	360	1 R80758	Interleukin 8 rece
30	10	2.8	360	1 R80953	Recombinant high a
31	10	2.8	355	1 R80950	Recombinant high a
32	10	2.8	350	1 R80951	Recombinant high a
33	10	2.8	358	1 R80952	Recombinant high a
34	10	2.8	367	1 W25217	Rat orphanin FQ re
35	10	2.8	367	1 W26582	Rat methadone-spec
36	10	2.8	73	1 W69988	Rodent chemokine r
37	10	2.8	110	1 W49091	Interleukin 8 type
38	10	2.8	367	1 W80549	Rat methadone-spec

ALIGNMENTS

RESULT 1
ID R68812 standard; Protein: 352 AA.
AC R68812:
DT 18-JUL-1995 (first entry)
DE Human monocyte PF4AR.
KW Interleukin-8 receptor; IL-8 receptor; PF4AR;
KW platelet factor superfamily receptor; monocyte; chemotactic;
KW inflammation; inflammatory disease; arthritis; emphysema; cystic;
KW fibrosis; colitis; bronchitis; meningitis; therapeutic.
OS Homo sapiens.
PN WO9428931-A.
PD 22-DEC-1994.
PF 07-JUN-1994: U06380.
PR 11-JUN-1993: US-076093.
PA (GETH) GENENTECH INC.
PI Chuntharapai A, Hebert C, Kim KJ, Lee J;
DR WPI: 95-036114/05.
DR N-PSDB: 080521.
PT Treatment of inflammatory disorders - by administering an
PT antibody capable of binding a platelet factor 4 superfamily
PI receptor polypeptide
PS Disclosure: page 54-56; 83pp; English.
CC 2 PF4AR members were identified by probing lambda libraries from
CC human monocyte-like cell line HL-60 and human peripheral blood
CC lymphocytes using a large fragment of IL-8 receptor DNA (full
CC sequence given in Q80520). The nucleotide sequences of the 2
CC PF4ARs are given in Q80521 and Q80522, and their respective
CC amino acid sequences in R68812 and R68813.
SQ Sequence 352 AA:

Query Match 100.0%; Score 352; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEGISITSDNYTEEMSGSDYSKMEPCFERENANFNIFPTIYSIFLTGIVNGIYI 60
DB 1 MEGISITSDNYTEEMSGSDYSKMEPCFERENANFNIFPTIYSIFLTGIVNGIYI 60
QY 61 LVMGQKRLRSTQDKYRLHLSTADLFLVITLPFAVDVANNYFENFLCKAVHYTYNL 120
DB 61 LVMGQKRLRSTQDKYRLHLSTADLFLVITLPFAVDVANNYFENFLCKAVHYTYNL 120
QY 121 YSSVITLAFISIDRLATVHATNSORPKLLAEKVYVYGVWIPALLTTPDIFANVSPA 180
DB 121 YSSVITLAFISIDRLATVHATNSORPKLLAEKVYVYGVWIPALLTTPDIFANVSPA 180
QY 181 DDRYICDRFPNDLWVVFQFQIHWGLIEPGIVLLSCYCIISKLSSHQGRKALKT 240
DB 181 DDRYICDRFPNDLWVVFQFQIHWGLIEPGIVLLSCYCIISKLSSHQGRKALKT 240
QY 241 TVIILAFACMLPRYIISTDSIFLLEIKQGCFFENTVHKWITITALAFHCLNPI 300
DB 241 TVIILAFACMLPRYIISTDSIFLLEIKQGCFFENTVHKWITITALAFHCLNPI 300
QY 301 LYAFGAKFTSAOHALTVSVSGSLKILSKGRGHSVSTESSSFHSS 352
DB 301 LYAFGAKFTSAOHALTVSVSGSLKILSKGRGHSVSTESSSFHSS 352
RESULT 2
ID R80757 standard; Protein: 352 AA.
AC R80757:
DT 26-MAR-1996 (first entry)
DE Chemokine superfamily receptor.
KW Chemokine; IL-8; inflammation; psoriasis; dermatitis;
KW rheumatoid arthritis; inflammatory bowel disease;
KW chronic lung inflammation; treatment; antibody;
KW affinity purification; detection.
OS Homo sapiens.
PN US5440021-A.

PD 08-AUG-1995.
 PF 29-MAR-1991: 677211.
 PR 29-MAR-1991: US-677211.
 PR 25-FEB-1994: US-202056.
 PR (CHUN/) CHUNTHARAPAI A.
 PA (HEBE/) HEBERT C.
 PA (KIM/) KIM K J.
 PA (LEEJ/) LEE J.
 PI Chuntharapai A, Hebert C, Kim KJ, Lee J:
 DR MPI: 95-283151/37.
 DR N-PSDB: 099007.
 PT New antibodies against Interleukin 8 type B receptor - used to treat
 or prevent inflammation, also for detecting receptor expression and
 purification.
 PS Example 2: Columns 45-48; 62pp; English.
 CC Antibodies directed against the interleukin-8 receptor B can be used
 to treat or prevent inflammation e.g. psoriasis, dermatitis,
 rheumatoid arthritis and particularly inflammatory bowel disease and
 chronic lung inflammation. When immunoblised, these antibodies may
 be used to detect Interleukin-8 receptor B expression in cells and
 tissues and for affinity purification of interleukin-8 receptor B
 from cells. This sequence is an additional chemokine superfamily
 receptor which was identified by probing lambda libraries of genomic
 DNA from a human monocyte-like cell line (U-60) and human peripheral
 blood lymphocytes using a large fragment of the interleukin-8 type
 A receptor DNA (See Q99006).
 CC Sequence 352 AA:
 SQ

Query Match 100.0%; Score 352; DB 1; Length 352;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEGISLYTSDNTTEENGSGDYDSMKPCFRENNANFNKIFLPTIYIIFLTGIVNGVLT 60
 DB 1 MEGISLYTSDNTTEENGSGDYDSMKPCFRENNANFNKIFLPTIYIIFLTGIVNGVLT 60

QY 61 LVMGYOKRIKRSMTDKRLHSVADLLFVITLPFAVDANVMYGFNLCRAVAVITVNL 120
 DB 61 LVMGYOKRIKRSMTDKRLHSVADLLFVITLPFAVDANVMYGFNLCRAVAVITVNL 120

QY 121 YSSVLLAFISLDRIYAIYHATNSOPRKLAEKVYVGVWIPALLITIPDIFANVSEA 180
 DB 121 YSSVLLAFISLDRIYAIYHATNSOPRKLAEKVYVGVWIPALLITIPDIFANVSEA 180

QY 181 DDEYICDREYNDLWVYVFOHIMWGLIPGIYILSCYIIISKLSHGOKRAKLT 240
 DB 181 DDEYICDREYNDLWVYVFOHIMWGLIPGIYILSCYIIISKLSHGOKRAKLT 240

QY 241 TVLLILAFACWLPYIYIGISIDSFILLEIKOGCEFNTHKMWISIEALAFHCCINPI 300
 DB 241 TVLLILAFACWLPYIYIGISIDSFILLEIKOGCEFNTHKMWISIEALAFHCCINPI 300

QY 301 LVAFILGAKFTSAQHALLTSVRSGLSLKILSKRGHSSVSTESSSSPHSS 352
 DB 301 LVAFILGAKFTSAQHALLTSVRSGLSLKILSKRGHSSVSTESSSSPHSS 352

RESULT 3
 W97362
 ID W97362 standard; Protein: 356 AA.
 AC W97362:
 DE 12-MAY-1999 (first entry) polypeptide designated CXCR4B.
 KW G-protein coupled receptor; CXCR4B; human; splice variant;
 KW chemokine receptor; CXCR4; vaccine; gene therapy; HIV-1; HIV-1;
 KW cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;
 KW parkinson's disease; acute heart failure; hypotension; hypertension;
 KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
 KW stroke; ulcer; allergy; benign prostatic hypertrophy; migraine;
 KW vomiting; psychotic disorder; neurological disorder; anxiety;
 KW schizophrenia; manic depression; delirium; dementia; mental retardation;
 KW dyskinesias; Huntingtons disease; Gilles de la Tourette syndrome.

OS Homo sapiens.
 PN EP-897980-A2.
 PD 24-FEB-1999.
 PF 07-AUG-1998; 306324.
 PR 24-JUL-1998; US-036601.
 PR 20-AUG-1997; US-056601.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Gupta SK, Pillariseti K;
 DR MPI: 99-134643/12.
 DR N-PSDB: X15882.
 PT New G protein coupled receptor (CXCR4B) polypeptide and
 polynucleotide, human splice variant of a chemokine receptor -
 useful as diagnostic reagents and for prevention and treatment of
 HIV infection, cancer, stroke and dementia
 PT Claim 1: Page 18; 24pp; English.

PS The present sequence represents a G-protein coupled receptor polypeptide
 designated CXCR4B, which is a human splice variant of a chemokine
 receptor CXCR4. CXCR4B polynucleotides and polypeptides are useful for
 diagnosing susceptibility to diseases by detecting mutations or
 polymorphisms in the CXCR4B gene or analysing for the presence or amount
 of CXCR4B polypeptide expressed in a patient. CXCR4B polypeptides and
 polynucleotides are also useful for screening for antagonists and
 agonists which can be used to treat conditions associated with CXCR4B
 agonists imbalance. CXCR4B polypeptides may be administered directly
 (as a vaccine) or via a vector (gene therapy) to prevent disease.
 CC Diseases diagnosed, prevented and treated include: HIV-1 and HIV-1
 infections; cancer; pain; diabetes; obesity; anorexia; bulimia; asthma;
 CC urinary retention; osteoporosis; angina pectoris; myocardial infarction;
 CC parkinson's disease; acute heart failure; hypotension; hypertension;
 CC stroke; ulcers; allergies; benign prostatic hypertrophy; migraine;
 CC vomiting; psychotic and neurological disorders including anxiety;
 CC schizophrenia, manic depression, delirium, dementia and severe mental
 CC retardation; and dyskinesias, such as Huntingtons disease or Gilles de
 CC la Tourette syndrome. CXCR4B polypeptides are useful for mapping genes
 CC to chromosomes, allowing gene inheritance to be studied through linkage
 CC analysis.
 CC Sequence 356 AA:
 SQ

Query Match 98.6%; Score 347; DB 1; Length 356;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 IYTSNDYTEMGSGDYDSMKPCFRENNANFNKIFLPTIYIIFLTGIVNGVLTVMGY 65
 DB 10 IYTSNDYTEMGSGDYDSMKPCFRENNANFNKIFLPTIYIIFLTGIVNGVLTVMGY 69

QY 66 OKRLSMTDKRYRLHSVADLLFVITLPFAVDANVMYGFNLCRAVAVITVNLXSVL 125
 DB 70 OKRLSMTDKRYRLHSVADLLFVITLPFAVDANVMYGFNLCRAVAVITVNLXSVL 129

QY 126 ILAFISLDRIYAIYHATNSOPRKLAEKVYVGVWIPALLITIPDIFANVSEADRYI 185
 DB 130 ILAFISLDRIYAIYHATNSOPRKLAEKVYVGVWIPALLITIPDIFANVSEADRYI 189

QY 186 CDRFYNDLWVYVFOHIMWGLIPGIYILSCYIIISKLSHGOKRAKLTIVILI 245
 DB 190 CDRFYNDLWVYVFOHIMWGLIPGIYILSCYIIISKLSHGOKRAKLTIVILI 249

QY 246 LAFACWLPYIYIGISIDSFILLEIKOGCEFNTHKMWISIEALAFHCCINPIYAF 305
 DB 250 LAFACWLPYIYIGISIDSFILLEIKOGCEFNTHKMWISIEALAFHCCINPIYAF 309

QY 306 GAKFTSAQHALLTSVRSGLSLKILSKRGHSSVSTESSSSPHSS 352
 DB 310 GAKFTSAQHALLTSVRSGLSLKILSKRGHSSVSTESSSSPHSS 356

RESULT 4
 W48734
 ID W48734 standard; Protein: 337 AA.
 AC W48734:
 DE 25-SEP-1998 (first entry)

DE Human RM3 seven transmembrane (7TM) receptor.
KW V28: placenta; seven transmembrane receptor; 7TM; signal transduction;
OS Immunology; Inflammation; RM3.
OS Homo sapiens.

FT Domain Location/Qualifiers
FT 48..69
FT /note= "Transmembrane domain 1"
FT 82..100
FT /note= "Transmembrane domain 2"
FT 115..136
FT /note= "Transmembrane domain 3"
FT 159..179
FT /note= "Transmembrane domain 4"
FT 198..220
FT /note= "Transmembrane domain 5"
FT 245..274
FT /note= "Transmembrane domain 6"
FT 287..311
FT /note= "Transmembrane domain 7"

PN U55759804-A.
PD 02-JUN-1998.
PF 17-NOV-1993: 153848.
PR 17-NOV-1992: US-977452.
PA (ICOS-) ICOS CORP.
PI Godiska R, Gray PM, Schweickart VL;
DR WPI: 98-332132/29.
DR N-PSDB: V18357.

PT DNA encoding V28 seven transmembrane receptor polypeptide - useful
PT for producing recombinant polypeptide and anti-V28 antibodies, and
PT in screening assays for V28 agonists and antagonists
PS Example 11: Columns 93-94: 56pp: English.

CC The present sequence represents the human RM3 protein which is encoded
CC by the full length genomic RM3 clone (V18357) which was isolated from
CC a human macrophage cDNA library. The invention claims for the full
CC length V28 genomic DNA sequence (V18343) isolated from a human placenta
CC genomic library. The V28 (M48722) and RM3 proteins are seven
CC transmembrane (7TM) receptors which are probably involved in signal
CC transduction. The invention also claims that cells transformed with V28
CC DNA can be used to produce the recombinant polypeptide, to produce
CC anti-V28 antibodies or in screening assays for V28 agonists or
CC antagonists. The antibodies, agonists and antagonists could then be
CC used to modulate V28 receptor-ligand binding, for e.g. in immunological
CC and/or inflammatory events in vivo.

SQ Sequence 337 AA:

Query Match 95.7%; Score 337; DB 1; Length 337;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 MGSDDYDSMKPCFREENANFNKIFLPTIYSIFLTGIVGGLVILVWGQKLRSMIDK 75
DB 1 MGSDDYDSMKPCFREENANFNKIFLPTIYSIFLTGIVGGLVILVWGQKLRSMIDK 60
OY 76 YRLHLSVADLLFVITLPFAVDAVANMYFGNFLCKAVHYITVNLSSVLLAFISLDYR 135
DB 61 YRLHLSVADLLFVITLPFAVDAVANMYFGNFLCKAVHYITVNLSSVLLAFISLDYR 120
OY 136 LAIYHAINSORPRKLLAEKVVYGVWIPALLITPDIFFANVSEADDRYICDRFPNDLW 195
DB 121 LAIYHAINSORPRKLLAEKVVYGVWIPALLITPDIFFANVSEADDRYICDRFPNDLW 180
OY 196 VVWFQFQIMWGLIPGIVILSCYCIISKLSHSGHOKRKALKTTVILIAFFACWLPY 255
DB 181 VVWFQFQIMWGLIPGIVILSCYCIISKLSHSGHOKRKALKTTVILIAFFACWLPY 240
OY 256 YIGSIDSFILLEIKKOCCEPENTYHKWISTEALAFHCCINPLIYVFLGAKERTSAQH 315
DB 241 YIGSIDSFILLEIKKOCCEPENTYHKWISTEALAFHCCINPLIYVFLGAKERTSAQH 300
OY 316 ALTSVRSGLSKILSKGRGSHSVSTESSESSPHSS 352
DB 301 ALTSVRSGLSKILSKGRGSHSVSTESSESSPHSS 337

RESULT 5
R27792 ID R27792 standard; Protein: 352 AA.
AC R27792:
DT 12-MAR-1993 (first entry)
DE New platelet factor 4 receptor superfamily member PF4AR1.
KW IL-8R; G-protein coupled receptor family; rhodopsin superfamily;
KW pro-inflammatory cytokine; 8tr.20.15.
OS Homo sapiens.
PN W09217497-A.
PD 15-OCT-1992.
PF 23-MAR-1992: U02317.
PR 29-MAR-1991: US-677211.
PR 19-DEC-1991: US-810782.
PA (GERTH) GENENTECH INC.
PI Holmes WE, Lee J, Wood WI;
DR WPI: 92-366191/44.
DR N-PSDB: Q29506.

PT Isolated human platelet factor 4 super-family receptor
PT polypeptide and corresp. antibodies and DNA - useful as
PT diagnostic and screening agents, and for treating inflammation or
PT PF4AR-mediated disorders
PS Claim 7: Fig 4: 78pp: English.

CC The IL-8 receptor cDNA sequence was isolated (see Q29505) and a
CC 874bp sub-fragment of the coding sequence was used as a probe to
CC screen human cell line HL60 and human peripheral blood lymphocyte
CC cDNA libraries. Two new gene sequences were found that are clearly
CC related to the IL-8 receptor. One of these was contained in
CC combined clone 8tr.20.15 and is predicted to encode an amino acid
CC sequence which is 34% identical with both the high and low affinity
CC IL-8 receptors. See also Q37107.

SQ Sequence 352 AA:

Query Match 77.6%; Score 273; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 1e-256;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MEGISITYSDNTEEMSGSDYDSMKPCFREENANFNKIFLPTIYSIFLTGIVGGLVI 60
DB 1 MEGISITYSDNTEEMSGSDYDSMKPCFREENANFNKIFLPTIYSIFLTGIVGGLVI 60
OY 61 LVMGQKRLRSTDKYRHLHLSVADLLFVITLPFAVDAVANMYFGNFLCKAVHYITVNL 120
DB 61 LVMGQKRLRSTDKYRHLHLSVADLLFVITLPFAVDAVANMYFGNFLCKAVHYITVNL 120
OY 121 YSSVILIAFISLDRLALVHATNSORPRKLLAEKVVYGVWIPALLITPDIFFANVSEA 180
DB 121 YSSVILIAFISLDRLALVHATNSORPRKLLAEKVVYGVWIPALLITPDIFFANVSEA 180
OY 181 DDRYICDRFPNDLWVWFQFQIMWGLIPGIVILSCYCIISKLSHSGHOKRKALKT 240
DB 181 DDRYICDRFPNDLWVWFQFQIMWGLIPGIVILSCYCIISKLSHSGHOKRKALKT 240
OY 241 TVIILAFACWLPYIIGSIDSFILLEIKKOG 273
DB 241 TVIILAFACWLPYIIGSIDSFILLEIKKOG 273

RESULT 6
R53753 ID R53753 standard; Protein: 337 AA.
AC R53753:
DT 07-FEB-1995 (first entry)
DE Seven transmembrane receptor (RM3).
KW Primer: seven transmembrane receptor; receptor; amplification; PCR;
KW polymerase chain reaction.
OS Homo sapiens.

FT Domain Location/Qualifiers
FT 48..69
FT /label= "Transmembrane domain".

PA (SHIO) SHIONOGI & CO LTD.
PI Iizasa H, Kishimoto T, Nagasawa T, Nakajima T, Tachibana K,
PI Yoshida N, Yoshie O;
DR MPI: 98-447232/38.
DR N-PSDB: V46370.
PT Mouse CXC chemokine receptor binding to PSF/SDF-1 pre-B cell
PT stimulatory factor - is useful for screening of potential HIV
PT infection and AIDS inhibitors
PS Claim 1: Pages 54-56: 75pp: Japanese.
CC The present sequence represents a murine CXC chemokine receptor which
CC binds to the mouse CXC chemokine pre-B cell stimulatory factor
CC PSF/SDF-1. The nucleic acid is isolated from mouse pre-B cell
CC line DM34. The receptor and cells expressing it can be used in the
CC study and mapping of the mechanism of HIV infection and in screening
CC of potential inhibitors of HIV infection and the development of AIDS.
SQ Sequence 359 AA:

Query Match 17.0%; Score 60; DB 1: Length 359;
Best Local Similarity 100.0%; Pred. No. 2.7e-50;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 VWFQFHIMWGLIPGIVLSCYCIISKLSHSGHQRKALKTTVILILAFACWLPY 256
|||||
Db 204 VWFQFHIMWGLIPGIVLSCYCIISKLSHSGHQRKALKTTVILILAFACWLPY 263

RESULT 9
W34001
ID W34001 standard: peptide: 25 AA.
AC W34001:
DT 28-MAY-1998 (first entry)
DE Fusin receptor antagonist.
KW Antagonist: integral membrane protein; inhibitor: transmembrane domain;
KW therapy: psychotic disorder; Huntington's disease; Parkinson's disease;
KW post-myocardial infarction; HIV infection; cell proliferation disorder;
KW peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease;
KW antibacterial; fusin receptor.
OS Homo sapiens.
PN WO9735881-A2.
PD 02-OCT-1997.
PF 26-MAR-1997; CA0203.
PR 20-AUG-1996; US-024240.
PR 27-MAR-1996; US-014306.
PR 25-JUN-1996; US-670119.
PA (GEOR/) GEORGE S R.
PA (NGCY/) NG G Y K.
PA (ODOM/) O'DOWD B F.
PA (SEEM/) SEEMAN P.
PI George SR, Ng GYK, O'Dowd BF, Seeman P;
DR MPI: 97-489566/45.
PT Inhibitor of integral membrane protein - used to treat e.g.
PT schizophrenia, hypertension, viral infection, cancer etc.
PS Disclosure: Page 88; 127pp; English.
CC This sequence represents a fusin receptor antagonist, and is an
CC antagonist of the invention. The antagonists (A) are for inhibiting the
CC function of a prokaryotic or eukaryotic integral membrane protein (IMP)
CC having at least one transmembrane domain (TMD), comprises a peptide
CC including at least 4 consecutive amino acids (aa) from the sequence of
CC the TMD. (A) are used to treat or prevent disorders in mammals that
CC involve disturbances of IMP, and the same effect is achieved by inserting
CC a nucleic acid encoding (A) into the cells of the mammal. Specified
CC conditions that can be treated are schizophrenia, psychotic disorders,
CC Huntington's disease, Tourette's syndrome, hypertension, post-myocardial
CC infarction, tachyarrhythmia, human immunodeficiency virus infection,
CC AIDS, cell proliferation (cancer), psoriasis or hyperkeratosis), and drug
CC abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma,
CC renal disease, Parkinson's disease, congestive heart failure, migraine,
CC depression, obesity, diabetic neuropathy/retinopathy, autoimmune
CC diseases, graft rejection etc. (A) can be used as diuretics and
CC antibacterials. (A) may also be used to determine function of
CC orphan receptors and, when attached to a marker, for tissue imaging to
CC localise/quantify specific receptors. The transgenic animals are used to

CC test efficacy/toxicity of (A) and as models of diseases. (A) are very
CC specific and selective, against the IMP from which they are derived, so
CC should have few side effects.
SQ Sequence 25 AA:

Query Match 7.1%; Score 25; DB 1: Length 25;
Best Local Similarity 100.0%; Pred. No. 2.1e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 FLPTIYSIFLTGIVNGVILVWG 64
|||||
Db 1 FLPTIYSIFLTGIVNGVILVWG 25

RESULT 10
W34003
ID W34003 standard: peptide: 22 AA.
AC W34003:
DT 28-MAY-1998 (first entry)
DE Fusin receptor antagonist.
KW Antagonist: integral membrane protein; inhibitor: transmembrane domain;
KW therapy: psychotic disorder; Huntington's disease; Parkinson's disease;
KW post-myocardial infarction; HIV infection; cell proliferation disorder;
KW peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease;
KW antibacterial; fusin receptor.
OS Homo sapiens.
PN WO9735881-A2.
PD 02-OCT-1997.
PF 26-MAR-1997; CA0203.
PR 20-AUG-1996; US-024240.
PR 27-MAR-1996; US-014306.
PR 25-JUN-1996; US-670119.
PA (GEOR/) GEORGE S R.
PA (NGCY/) NG G Y K.
PA (ODOM/) O'DOWD B F.
PA (SEEM/) SEEMAN P.
PI George SR, Ng GYK, O'Dowd BF, Seeman P;
DR MPI: 97-489566/45.
PT Inhibitor of integral membrane protein - used to treat e.g.
PT schizophrenia, hypertension, viral infection, cancer etc.
PS Disclosure: Page 88; 127pp; English.
CC This sequence represents a fusin receptor antagonist, and is an
CC antagonist of the invention. The antagonists (A) are for inhibiting the
CC function of a prokaryotic or eukaryotic integral membrane protein (IMP)
CC having at least one transmembrane domain (TMD), comprises a peptide
CC including at least 4 consecutive amino acids (aa) from the sequence of
CC the TMD. (A) are used to treat or prevent disorders in mammals that
CC involve disturbances of IMP, and the same effect is achieved by inserting
CC a nucleic acid encoding (A) into the cells of the mammal. Specified
CC conditions that can be treated are schizophrenia, psychotic disorders,
CC Huntington's disease, Tourette's syndrome, hypertension, post-myocardial
CC infarction, tachyarrhythmia, human immunodeficiency virus infection,
CC AIDS, cell proliferation (cancer), psoriasis or hyperkeratosis), and drug
CC abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma,
CC renal disease, Parkinson's disease, congestive heart failure, migraine,
CC depression, obesity, diabetic neuropathy/retinopathy, autoimmune
CC diseases, graft rejection etc. (A) can be used as diuretics and
CC antibacterials. (A) may also be used to determine function of
CC orphan receptors and, when attached to a marker, for tissue imaging to
CC localise/quantify specific receptors. The transgenic animals are used to
CC test efficacy/toxicity of (A) and as models of diseases. (A) are very
CC specific and selective, against the IMP from which they are derived, so
CC should have few side effects.
SQ Sequence 22 AA:

Query Match 6.2%; Score 22; DB 1: Length 22;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 AVHYITVNIYSSVLIAPISL 132
|||||

Db 1 AVHVIYTNLYSSVILAFISL 22

RESULT 11
W34005 standard: peptide: 22 AA.

AC W34005: (first entry)
DE 28-MAY-1998 (first entry)
Fusin receptor antagonist.
KW Antagonist; integral membrane protein; inhibitor; transmembrane domain; therapy; psychotic disorder; Huntington's disease; Parkinson's disease; post-myocardial infarction; HIV infection; cell proliferation disorder; peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease; antibacterial; fusin receptor.
OS Homo sapiens.
PN WO9735881-A2.
PD 02-OCT-1997.
PE 26-MAR-1997; CA0203.
PR 20-AUG-1996; US-024240.
PR 27-MAR-1996; US-014306.
PR 25-JUN-1996; US-670119.
PA (GEORGE) GEORGE S R.
PA (NGUY) NG G Y K.
PA (ODOM/) O'DOWD B F.
PA (SEEM/) SEEMAN P.
PI George SR, Ng GYK, O'Dowd BF, Seeman P; WPI: 97-489566/45.
PT Inhibitor of integral membrane protein - used to treat e.g. schizophrenia, hypertension, viral infection, cancer etc.
PS Disclosure: Page 88; 127pp; English.
CC This sequence represents a fusin receptor antagonist, and is an antagonist of the invention. The antagonists (A) are for inhibiting the function of a prokaryotic or eukaryotic integral membrane protein (IMP) having at least one transmembrane domain (TMD), comprises a peptide including at least 4 consecutive amino acids (aa) from the sequence of the TMD. (A) are used to treat or prevent disorders in mammals that involve disturbances of IMP, and the same effect is achieved by inserting a nucleic acid encoding (A) into the cells of the mammal. Specified conditions that can be treated are schizophrenia, psychotic disorders, Huntington's disease, Tourette's syndrome, hypertension, post-myocardial infarction, tachyarrhythmia, human immunodeficiency virus infection, AIDS, cell proliferation (cancer, psoriasis or hyperkeratosis), and drug abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma, renal disease, Parkinson's disease, congestive heart failure, migraine, depression, obesity, diabetic neuropathy/retinopathy, autoimmune diseases, graft rejection etc. (A) can be used as diuretics and antibacterials. (A) may also be used to determine function of orphan receptors and, when attached to a marker, for tissue imaging to localise/quantify specific receptors. The transgenic animals are used to test efficacy/toxicity of (A) and as models of diseases. (A) are very specific and selective against the IMP from which they are derived, so CC should have few side effects.
SQ Sequence 22 AA:

Query Match 6.2%; Score 22; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 22: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 201 FOHIMWGILPGIVILSCYCI 222
|||||
Db 1 FOHIMWGILPGIVILSCYCI 22

RESULT 12
W34004 standard: peptide: 21 AA.
AC W34004:
DE 28-MAY-1998 (first entry)
Fusin receptor antagonist.
KW Antagonist; integral membrane protein; inhibitor; transmembrane domain; therapy; psychotic disorder; Huntington's disease; Parkinson's disease; post-myocardial infarction; HIV infection; cell proliferation disorder;

KW peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease; antibacterial; fusin receptor.

OS Homo sapiens.
PN WO9735881-A2.
PD 02-OCT-1997.
PE 26-MAR-1997; CA0203.
PR 20-AUG-1996; US-024240.
PR 27-MAR-1996; US-014306.
PR 25-JUN-1996; US-670119.
PA (GEORGE) GEORGE S R.
PA (NGUY) NG G Y K.
PA (ODOM/) O'DOWD B F.
PA (SEEM/) SEEMAN P.
PI George SR, Ng GYK, O'Dowd BF, Seeman P; WPI: 97-489566/45.
PT Inhibitor of integral membrane protein - used to treat e.g. schizophrenia, hypertension, viral infection, cancer etc.
PS Disclosure: Page 88; 127pp; English.
CC This sequence represents a fusin receptor antagonist, and is an antagonist of the invention. The antagonists (A) are for inhibiting the function of a prokaryotic or eukaryotic integral membrane protein (IMP) having at least one transmembrane domain (TMD), comprises a peptide including at least 4 consecutive amino acids (aa) from the sequence of the TMD. (A) are used to treat or prevent disorders in mammals that involve disturbances of IMP, and the same effect is achieved by inserting a nucleic acid encoding (A) into the cells of the mammal. Specified conditions that can be treated are schizophrenia, psychotic disorders, Huntington's disease, Tourette's syndrome, hypertension, post-myocardial infarction, tachyarrhythmia, human immunodeficiency virus infection, AIDS, cell proliferation (cancer, psoriasis or hyperkeratosis), and drug abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma, renal disease, Parkinson's disease, congestive heart failure, migraine, depression, obesity, diabetic neuropathy/retinopathy, autoimmune diseases, graft rejection etc. (A) can be used as diuretics and antibacterials. (A) may also be used to determine function of orphan receptors and, when attached to a marker, for tissue imaging to localise/quantify specific receptors. The transgenic animals are used to test efficacy/toxicity of (A) and as models of diseases. (A) are very specific and selective against the IMP from which they are derived, so CC should have few side effects.
SQ Sequence 21 AA:

Query Match 6.0%; Score 21; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 155 VYVGWIPALLITPPIFA 175
|||||
Db 1 VYVGWIPALLITPPIFA 21

RESULT 13
W34006 standard: peptide: 21 AA.

AC W34006: (first entry)
DE 28-MAY-1998 (first entry)
Fusin receptor antagonist.
KW Antagonist; integral membrane protein; inhibitor; transmembrane domain; therapy; psychotic disorder; Huntington's disease; Parkinson's disease; post-myocardial infarction; HIV infection; cell proliferation disorder; peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease; antibacterial; fusin receptor.
OS Homo sapiens.
PN WO9735881-A2.
PD 02-OCT-1997.
PE 26-MAR-1997; CA0203.
PR 20-AUG-1996; US-024240.
PR 27-MAR-1996; US-014306.
PR 25-JUN-1996; US-670119.
PA (GEORGE) GEORGE S R.
PA (NGUY) NG G Y K.
PA (ODOM/) O'DOWD B F.

PA (SEEM/) SEEMAN P.
 PI George SR, NG GYK, ODOwd BF, Seeman P;
 DR WPI: 97-489566/45.
 CC Inhibitor of integral membrane protein - used to treat e.g.
 PT schizophrenia, hypertension, viral infection, cancer etc.
 PS Disclosure: Page 88; 127pp; English.
 CC This sequence represents a fusin receptor antagonist, and is an
 CC antagonist of the invention. The antagonists (A) are for inhibiting the
 CC function of a prokaryotic or eukaryotic integral membrane protein (IMP)
 CC having at least one transmembrane domain (TMD), comprises a peptide
 CC including at least 4 consecutive amino acids (aa) from the sequence of
 CC the TMD. (A) are used to treat or prevent disorders in mammals that
 CC involve disturbances of IMP, and the same effect is achieved by inserting
 CC a nucleic acid encoding (A) into the cells of the mammal. Specified
 CC conditions that can be treated are schizophrenia, psychotic disorders,
 CC Huntington's disease, Tourette's syndrome, hypertension, post-myocardial
 CC infarction, tachyarrhythmia, human immunodeficiency virus infection,
 CC AIDS, cell proliferation (cancer, psoriasis or hyperkeratosis), and drug
 CC abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma,
 CC renal disease, Parkinson's disease, congestive heart failure, migraine,
 CC depression, obesity, diabetic neuropathy/retinopathy, autoimmune
 CC diseases, graft rejection etc. (A) can be used as diuretics and
 CC antibacterials. (A) may also be used to determine function of
 CC orphan receptors and, when attached to a marker, for tissue imaging to
 CC localise/quantify specific receptors. The transgenic animals are used to
 CC test efficacy/toxicity of (A) and as models of diseases. (A) are very
 CC specific and selective, against the IMP from which they are derived, so
 CC should have few side effects.
 SQ Sequence 21 AA:

Query Match 5.0%; Score 21; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1,3e-13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 241 TVILIAFFACWLPIYIGISI 261
 ||||||||||||||||||||
 Db 1 TVILIAFFACWLPIYIGISI 21

RESULT 14
 W34007
 ID W34007 standard; peptide: 21 AA.
 AC W34007:
 DT 28-MAY-1998 (first entry)
 DE Fusin receptor antagonist.
 KW Antagonist; integral membrane protein; inhibitor; transmembrane domain;
 KW therapy; psychotic disorder; Huntington's disease; Parkinson's disease;
 KW post-myocardial infarction; HIV infection; cell proliferation disorder;
 KW peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease;
 KW antibacterial; fusin receptor.
 OS Homo sapiens.
 PN WO9735881-A2.
 PD 02-OCT-1997.
 PF 26-MAR-1997; CA0203.
 PR 20-AUG-1996; US-024240.
 PR 27-MAR-1996; US-014306.
 PR 25-JUN-1996; US-670119.
 PA (GEOR/) GEORGE S R.
 PA (NGCY/) NG G Y K.
 PA (ODOW/) O'DOWD B F.
 PA (SEEM/) SEEMAN P.
 PI George SR, NG GYK, ODOwd BF, Seeman P;
 DR WPI: 97-489566/45.
 CC Inhibitor of integral membrane protein - used to treat e.g.
 PT schizophrenia, hypertension, viral infection, cancer etc.
 PS Disclosure: Page 88; 127pp; English.
 CC This sequence represents a fusin receptor antagonist, and is an
 CC antagonist of the invention. The antagonists (A) are for inhibiting the
 CC function of a prokaryotic or eukaryotic integral membrane protein (IMP)
 CC having at least one transmembrane domain (TMD), comprises a peptide
 CC including at least 4 consecutive amino acids (aa) from the sequence of
 CC the TMD. (A) are used to treat or prevent disorders in mammals that

CC involve disturbances of IMP, and the same effect is achieved by inserting
 CC a nucleic acid encoding (A) into the cells of the mammal. Specified
 CC conditions that can be treated are schizophrenia, psychotic disorders,
 CC Huntington's disease, Tourette's syndrome, hypertension, post-myocardial
 CC infarction, tachyarrhythmia, human immunodeficiency virus infection,
 CC AIDS, cell proliferation (cancer, psoriasis or hyperkeratosis), and drug
 CC abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma,
 CC renal disease, Parkinson's disease, congestive heart failure, migraine,
 CC depression, obesity, diabetic neuropathy/retinopathy, autoimmune
 CC diseases, graft rejection etc. (A) can be used as diuretics and
 CC antibacterials. (A) may also be used to determine function of
 CC orphan receptors and, when attached to a marker, for tissue imaging to
 CC localise/quantify specific receptors. The transgenic animals are used to
 CC test efficacy/toxicity of (A) and as models of diseases. (A) are very
 CC specific and selective, against the IMP from which they are derived, so
 CC should have few side effects.
 SQ Sequence 21 AA:

Query Match 5.0%; Score 21; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1,3e-13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 286 ITEALAFHCCLNPIYAFLG 306
 ||||||||||||||||||||
 Db 1 ITEALAFHCCLNPIYAFLG 21

RESULT 15
 W34002
 ID W34002 standard; peptide: 20 AA.
 AC W34002:
 DT 28-MAY-1998 (first entry)
 DE Fusin receptor antagonist.
 KW Antagonist; integral membrane protein; inhibitor; transmembrane domain;
 KW therapy; psychotic disorder; Huntington's disease; Parkinson's disease;
 KW post-myocardial infarction; HIV infection; cell proliferation disorder;
 KW peptic ulcer; glaucoma; asthma; renal disease; autoimmune disease;
 KW antibacterial; fusin receptor.
 OS Homo sapiens.
 PN WO9735881-A2.
 PD 02-OCT-1997.
 PF 26-MAR-1997; CA0203.
 PR 20-AUG-1996; US-024240.
 PR 27-MAR-1996; US-014306.
 PR 25-JUN-1996; US-670119.
 PA (GEOR/) GEORGE S R.
 PA (NGCY/) NG G Y K.
 PA (ODOW/) O'DOWD B F.
 PA (SEEM/) SEEMAN P.
 PI George SR, NG GYK, ODOwd BF, Seeman P;
 DR WPI: 97-489566/45.
 CC Inhibitor of integral membrane protein - used to treat e.g.
 PT schizophrenia, hypertension, viral infection, cancer etc.
 PS Disclosure: Page 88; 127pp; English.
 CC This sequence represents a fusin receptor antagonist, and is an
 CC antagonist of the invention. The antagonists (A) are for inhibiting the
 CC function of a prokaryotic or eukaryotic integral membrane protein (IMP)
 CC having at least one transmembrane domain (TMD), comprises a peptide
 CC including at least 4 consecutive amino acids (aa) from the sequence of
 CC the TMD. (A) are used to treat or prevent disorders in mammals that
 CC involve disturbances of IMP, and the same effect is achieved by inserting
 CC a nucleic acid encoding (A) into the cells of the mammal. Specified
 CC conditions that can be treated are schizophrenia, psychotic disorders,
 CC Huntington's disease, Tourette's syndrome, hypertension, post-myocardial
 CC infarction, tachyarrhythmia, human immunodeficiency virus infection,
 CC AIDS, cell proliferation (cancer, psoriasis or hyperkeratosis), and drug
 CC abuse. (A) can also be used to treat peptic ulcer, glaucoma, asthma,
 CC renal disease, Parkinson's disease, congestive heart failure, migraine,
 CC depression, obesity, diabetic neuropathy/retinopathy, autoimmune
 CC diseases, graft rejection etc. (A) can be used as diuretics and
 CC antibacterials. (A) may also be used to determine function of
 CC orphan receptors and, when attached to a marker, for tissue imaging to

CC localise/quantify specific receptors. The transgenic animals are used to
 CC test efficacy/toxicity of (A) and as models of diseases. (A) are very
 CC specific and selective, against the IMP from which they are derived, so
 CC should have few side effects.
 SQ Sequence 20 AA:

Query Match 3.4%; Score 12; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 6.7e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 88 VITLPPMAVDAY 99
 |||||
 Db 9 VITLPPMAVDAY 20

RESULT 16

R27791 ID R27791 standard; Protein; 350 AA.

AC R27791; 12-MAR-1993 (first entry)
 DE Interleukin-8 receptor.
 KW IL-8R; G-protein coupled receptor family; rhodopsin superfamily;
 KW pro-inflammatory cytokine.

OS Homo sapiens. Location/Qualifiers
 FH Key 40..67
 FT domain /label= transmembrane
 FT /note= "putative"
 FT 76..98
 FT domain /label= transmembrane
 FT /note= "putative"
 FT 112..133
 FT domain /label= transmembrane
 FT /note= "putative"
 FT 155..174
 FT domain /label= transmembrane
 FT /note= "putative"
 FT 204..226
 FT domain /label= transmembrane
 FT /note= "putative"
 FT 243..264
 FT domain /label= transmembrane
 FT /note= "putative"
 FT 292..312
 FT domain /label= transmembrane
 FT /note= "putative"

PN WC9217497-A.
 PD 15-OCT-1992.
 PF 23-MAR-1992; U02317.
 PR 29-MAR-1991; US-677211.
 PR 19-DEC-1991; US-810782.
 PA (GETH) GENENTECH INC.
 PI Holmes WE, Lee J, Wood WI;
 WP1: 92-366191/44.
 DR N-PSDB: Q29505.
 PT Isolated human platelet factor 4 super-family receptor
 PT polypeptide and corresp. antibodies and DNA - useful as
 PT diagnostic and screening agents, and for treating inflammation or
 PT PF4AR-mediated disorders
 PS Claim 7; Fig 2; 78pp; English.
 CC A cDNA library constructed from human neutrophil mRNA in the mammalian
 CC expression vector pRK5B was transfected into COS-7 cells as pools of
 CC 2500 clones. One positive pool from the first 58 transfections was
 CC partitioned into smaller pools until a pure clone (pRK5B.116f1.1)
 CC was obtained. The ORF encodes a protein of 350 amino acids which
 CC shares several features with the G-protein coupled receptors of the
 CC rhodopsin superfamily, including 7 hydrophobic (transmembrane)
 CC domains. The IL-8 receptor is a preferred PF4AR superfamily member
 CC of the invention. See also Q29506 and Q37107.
 SQ Sequence 350 AA:

Query Match 2.8%; Score 10; DB 1; Length 350;
 Best Local Similarity 100.0%; Pred. No. 0.075;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 Db 134 DRYLAIVHAT 143

RESULT 17

R28272 ID R28272 standard; Protein; 355 AA.

AC R28272; 04-APR-1993 (first entry)
 DE Sequence in a high affinity recombinant rabbit interleukin-8
 DE (IL-8) receptor polypeptide in F3R.
 KW IL-8 receptor polypeptide; G-protein-coupled receptor.
 OS Oryctolagus cuniculus.

PN W09218641-A.
 PD 29-OCT-1992.
 PF 10-APR-1992; U02977.
 PR 10-APR-1991; US-685101.
 PR 09-JUL-1991; US-726606.
 PR 09-DEC-1991; US-803842.
 PA (REPR) REPLIGEN CORP.
 PA (UYBO-) UNIV BOSTON.
 PI Navarro J, Thomas KM, Wilt DP;
 WP1: 92-382123/46.
 DR N-PSDB: Q30011.
 PT Recombinant mammalian interleukin-8 receptor - used for screening
 PT Interleukin-8 binding antagonists, used to treat inflammation
 PS Claim 2; Fig 1; 71pp; English.
 CC Rabbit high affinity IL-8 receptor gene was isolated from rabbit
 CC peritoneal neutrophils and used as a source of poly(A)+ RNA, to
 CC produce a rabbit neutrophil cDNA library. 250,000 recombinant
 CC plaques were screened for those which hybridized to an antisense
 CC oligonucleotide (Q30015). This probe was designed based on the
 CC sequence derived from the second transmembrane domain of G-protein-
 CC coupled receptors. After tertiary screening, six plaques were
 CC isolated. The insert of one of these plaques, termed F3R was of 2.5
 CC kb in size. This insert was sequenced. The protein deduced from
 CC the F3R clone demonstrates that it belongs to the family of
 CC G-protein-coupled receptors. The deduced protein sequence
 CC indicates seven putative transmembrane segments.
 SQ Sequence 355 AA:

Query Match 2.8%; Score 10; DB 1; Length 355;
 Best Local Similarity 100.0%; Pred. No. 0.076;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 Db 139 DRYLAIVHAT 148

RESULT 18

R28273 ID R28273 standard; Protein; 360 AA.

AC R28273; 04-APR-1993 (first entry)
 DE Sequence in a low affinity recombinant human interleukin-8
 DE (IL-8) receptor polypeptide in 4AB.
 KW IL-8 receptor polypeptide; G-protein-coupled receptor.
 OS Homo sapiens.

PN W09218641-A.
 PD 29-OCT-1992.
 PF 10-APR-1992; U02977.
 PR 10-APR-1991; US-685101.
 PR 09-JUL-1991; US-726606.
 PR 09-DEC-1991; US-803842.
 PA (REPR) REPLIGEN CORP.
 PA (UYBO-) UNIV BOSTON.

PI Navarro J Thomas KM, Wilt DP;
 DR WPI: 92-362123/46.
 DR N-PSDB: Q30012.
 PT Recombinant mammalian interleukin-8 receptor - used for screening
 PT interleukin-8 binding antagonists, used to treat inflammation
 PS Disclosure: Fig 2: 71pp; English.
 CC Rabbit high affinity IL-8 receptor gene was isolated from rabbit
 CC peritoneal neutrophils and used as a source of poly(A)⁺ RNA, to
 CC produce a rabbit neutrophil cDNA library. 250,000 recombinant
 CC plaques were screened for those which hybridized to an antisense
 CC oligonucleotide (Q30015). This probe was designed based on the
 CC sequence derived from the second transmembrane domain of G-protein-
 CC coupled receptors. After tertiary screening, six plaques were
 CC isolated. The insert of one of these plaques, termed F3R was of 2.5
 CC kb in size. This insert was sequenced. The protein deduced from
 CC the F3R clone demonstrates that it belongs to the family of
 CC G-protein-coupled receptors. The deduced protein sequence
 CC indicates seven putative transmembrane segments. A human
 CC peripheral blood leukocyte lambda gtil cDNA library (5' stretch) was
 CC screened with a 652 bp EcoRI/BamHI fragment (including nucleotides
 CC -27 to 625) of the rabbit F3R clone. After tertiary screening
 CC several human clones which hybridized to the rabbit IL-8 probe
 CC were isolated. The insert of one such clone, termed 4AB was
 CC sequenced (Q30012).
 SO Sequence 360 AA;

Query Match 2.8%; Score 10; DB 1; Length 360;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 DB 143 DRYLAIVHAT 152

RESULT 19
 R33420
 ID R33420 standard; Protein: 355 AA.
 AC R33420:
 DT 26-JUL-1993 (first entry)
 DE Human IL-8 receptor from clone p2.
 KW Interleukin-8 receptor; probes; gene therapy; gro receptor;
 KW intracellular-calcium mobilising; ligand-binding; MIP-2 receptor.
 OS Homo sapiens.
 PN WO9306229-A.
 PD 01-APR-1993:
 PF 14-SEP-1992: U07641.
 PR 13-SEP-1991: US-759568.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Murphy PM:
 DR WPI: 93-117549/14.
 DR N-PSDB: Q38747.
 PT New interleukin-8 receptor aminoacid sequence - and corresp. cDNA
 PT expressed in Xenopus laevis oocytes or transfected host cells,
 PT for screening ligands of IL-8 receptor and gene therapy
 PS Claim 1: Fig 3: 39pp; English.
 CC cDNA libraries from 2 and 3.5 kb fractions of poly(A)⁺ RNA from HL60
 CC neutrophils sept. of a sucrose gradient were made in UniAmp. The
 CC libraries were screened with F3R oligonucleotide probe (from rabbit
 CC IL-8 receptor) and under low stringency with a p2 cDNA probe
 CC synthesised from random primers, to isolated the clone p2, encoding
 CC human IL-8 receptor protein. The presence or absence of the DNA
 CC coding IL-8R or related MIP-2 receptor may be detected using
 CC portions of the p2 clone as probes. p2 may also be used to screen
 CC for ligands of IL-8R and may also be used in gene therapy to treat a
 CC patient deficient in IL-8R. The IL-8R is a gro receptor and has
 CC intracellular calcium-mobilising and ligand-binding properties.
 SO Sequence 355 AA;

Query Match 2.8%; Score 10; DB 1; Length 355;
 Best Local Similarity 100.0%; Pred. No. 0.076;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 DB 138 DRYLAIVHAT 147

RESULT 20
 ID R37903 standard; peptide: 18 AA.
 AC R37903:
 DT 05-NOV-1993 (first entry)
 DE Interleukin-8 receptor second cytoplasmic loop analogue.
 KW Inflammation; inflammatory response; G-protein-coupled receptor;
 KW phagocyte; chemotactic gradient; anti-inflammatory; IL-8-R;
 KW chemotactic receptor; formylpeptide receptor; immunosuppression.
 OS Synthetic.
 PN WO9311784-A.
 PD 24-JUN-1993:
 PF 11-DEC-1992: U10580.
 PR 13-DEC-1991: US-608453.
 PA (UYDU-) UNIV DUKE.
 PI Didsbury JR, Snyderman R, Uhlig RJ;
 DR WPI: 93-213817/26.
 PT New antiinflammatory or pro-inflammatory peptide(s) - inhibit
 PT chemo-attractant receptor-protein G interaction or receptor
 PT desensitisation, for treating arthritis, psoriasis,
 PT immunosuppression, etc.
 PS Claim 59: Page 59: 76pp; English.
 CC Analogue peptides based on the sequence of chemotactic receptors
 CC regions in the second cytoplasmic loop of positively-charged
 CC can inhibit the inflammatory response. Peptides were designed based
 CC on regions of homology on the second cytoplasmic loop of known
 CC receptors. R37903 is based on a sequence from the interleukin-8
 CC receptor. Peptides having this sequence or having a sequence of
 CC not more than 50 amino acids and comprising this 18mer are preferred
 CC anti-inflammatory compounds of the invention.
 CC See R37901-R37904.
 SO Sequence 18 AA;

Query Match 2.8%; Score 10; DB 1; Length 18;
 Best Local Similarity 100.0%; Pred. No. 0.0053;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 DB 1 DRYLAIVHAT 10

RESULT 21
 ID R68811 standard; Protein: 350 AA.
 AC R68811:
 DT 18-JUL-1995 (first entry)
 DE Interleukin-8 receptor.
 KW Interleukin-8 receptor; IL-8 receptor; PF44:
 KW platelet factor superfamily receptor; neutrophil; chemotactic;
 KW inflammation; inflammatory disease; arthritis; emphysema; cystic;
 KW fibrosis; colitis; bronchitis; meningitis; therapeutic.
 OS Homo sapiens.
 PN WO9428931-A.
 PD 22-DEC-1994:
 PF 07-JUN-1994: U06380.
 PR 11-JUN-1993: US-076093.
 PA (GETH) GENENTECH INC.
 PI Chuntarapai A, Hebert C, Kim KJ, Lee J;
 DR WPI: 95-036114/05.
 DR N-PSDB: Q80520.
 PT Treatment of inflammatory disorders - by administering an
 PT antibody capable of binding a platelet factor 4 superfamily
 PT receptor polypeptide
 PS Disclosure: Page 51-54; 83pp; English.

CC A cDNA library constructed from human neutrophil mRNA in pRc5B was
 CC transfected into COS-7 cells, and the cells were screened with 125I-
 CC IL-8. The DNA sequence of isolated cDNA clone pRc5B.118r1.1,
 CC encoding human IL-8 receptor, is given in Q80520 and the predicted
 CC amino acid sequence in R68811. The receptor is used to raise
 CC antibodies that neutralize the activity of PF4AR, e.g. IL-8 receptor.
 SQ Sequence 350 AA;

Query Match 2.8%; Score 10; DB 1; Length 350;
 Best Local Similarity 100.0%; Pred. No. 0.075;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DRYLAIVHAT 142
 |||||
 DB 134 DRYLAIVHAT 143

RESULT 22

R67671 standard: Protein; 367 AA.
 AC R67671:
 DI 17-AUG-1995 (first entry)
 DE Mouse opiod receptor-like receptor MOP2 protein.
 KM Mouse; kappa; delta; mu; opiod receptor; brain; primer: PCR; amplify;
 KW Transmembrane domain; somatostatin; receptor; human; expression vector;
 KW truncate; chimeric; assay; probe.
 OS Mus musculus.
 PN WO9428132-A.
 PD 08-DEC-1994.
 PF 20-MAY-1994; 005747.
 PR 20-MAY-1993; US-066296.
 PR 30-JUL-1993; US-100694.
 PR 05-NOV-1993; US-147592.
 PA (ARCH-) ARCH DEV CORP.
 PI Bell GI, Reisine T, Yasuda K;
 DR WPI: 95-022804/03.
 DR N-PSDB: 075928.

PT Polynucleotides and peptides derived from opiod receptor
 PT screening assays for useful therapeutic compositions and in
 PT Polypeptides - for use in therapeutic substances.
 PS Claim 27; Page 225-229; 300pp; English.
 CC The amino acid sequence of the novel mouse opiod receptor-like receptor
 CC MOP2. MOP2 is a mouse receptor with pharmacological properties which
 CC are dissimilar to the properties of classic opiod receptors such kappa,
 CC delta, mu or sigma. It has been found that drugs of high abuse potential
 CC or analgesic potency bind selectively to this receptor. This suggests
 CC that this receptor could be important in the development of drugs to
 CC treat addiction. Other opiod receptors isolated and produced such as the
 CC novel mouse kappa and delta opiod receptors (Q75926-7) are useful for
 CC the development of novel assays designed to select or improve substances,
 CC capable of interacting with the opiod receptor proteins, for use in
 CC diagnosis, drug design and therapeutic applications.
 SQ Sequence 367 AA;

Query Match 2.8%; Score 10; DB 1; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 CLNPILYAFL 305
 |||||
 DB 310 CLNPILYAFL 319

RESULT 23

R71968 standard: Protein; 367 AA.
 AC R71968:
 DI 20-OCT-1995 (first entry)
 DE Rat opiod receptor.
 KW Opiod receptor; gene therapy; diagnostic.
 OS Rattus sp.

PN WO9507983-A.
 PD 23-MAR-1995.
 PF 13-SEP-1994; U10358.
 PR 13-SEP-1993; US-120601.
 PA (INDV) UNIV INDIANA FOUND.
 PI Yu L;
 DR WPI: 95-131351/17.
 DR N-PSDB: 089233.
 PT New nucleic acid encoding new human mu opiod receptor - and
 PT related vectors, transformed cells, antibodies etc., useful in
 PT diagnosis, treatment and drug screening.
 PS Example 9; Page 218-222; 266pp; English.
 CC The CDNA given in Q89233 was isolated from a rat brain library by
 CC low stringency hybridization with rat mu opiod receptor CDNA
 CC (Q89222). The clone encoded a 367-amino acid protein (R71968)
 CC that showed high homology with mu, kappa and delta opiod receptors
 CC but lacked affinity for their ligands, suggesting it to be
 CC a novel member of the opiod receptor family.
 SQ Sequence 367 AA;

Query Match 2.8%; Score 10; DB 1; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 CLNPILYAFL 305
 |||||
 DB 310 CLNPILYAFL 319

RESULT 24

R74298 standard: Protein; 367 AA.
 AC R74298:
 DI 03-NOV-1995 (first entry)
 DE Mouse kappa-3 opiod receptor.
 KW Kappa-3 opiod receptor; analgesia.
 OS Mus sp.
 PN WO9512616-A.
 PD 11-MAY-1995.
 PF 03-NOV-1994; U12728.
 PR 05-NOV-1993; US-147949.
 PA (SLOK) SLOAN KETERING INST CANCER RES.
 PI Pan Y, Pasternak GW;
 DR WPI: 95-193814/25.
 DR N-PSDB: 090096.
 PT Nucleic acid molecule(s) encoding a kappa-3 opiod receptor, and
 PT antibody against the receptor - used to detect the receptor, and to
 PT image cell membrane-bound receptor in a subject
 PS Disclosure; Fig.1; 68pp; English.
 CC Degenerate primers based on conserved sequences of the mouse delta
 CC opiod receptor were used in PCR to amplify mouse kappa-3 opiod
 CC receptor CDNA from a mouse brain lambda ZAP CDNA library. The
 CC encoded protein was expressed in E. coli and COS-7 cells.
 SQ Sequence 367 AA;

Query Match 2.8%; Score 10; DB 1; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 CLNPILYAFL 305
 |||||
 DB 310 CLNPILYAFL 319

RESULT 25

R76638 standard: Protein; 367 AA.
 AC R76638:
 DI 22-DEC-1995 (first entry)
 DE Rat opiorph receptor OR7.
 KW Opiorph receptor; opiod.

OS Rattus sp.
 PN W09519986-A1.
 PD 27-JUL-1995.
 PF 20-JAN-1995; U00939.
 PR 21-JAN-1994; US-185360.
 PA (AMCY) AMERICAN CYANAMID CO.
 PI Eppler CM, Hulmes JD, Ozenberger BA.
 DR WPI: 95-269412/35.
 DR N-PSDB: Q92972.
 PT New isolated DNA encoding an opiorph receptor - used to develop
 PT prods. for identifying opiorph agonists and antagonists and for
 PT detection and manipulation
 PS Disclosure: Fig.1: 35pp; English.
 CC Primers based on rodent opiorph receptors were used to amplify rat
 CC genomic DNA. Products were re-amplified and subcloned into PCR-II
 CC vector and amplified in E. coli. Plasmid DNAs were isolated, and
 CC an unspliced sequence was obtd. (Q92972) encoding the transmembrane
 CC domain opiorph receptor protein OR7.
 SO Sequence 367 AA:

Query Match 2.8%; Score 10; DB 1; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.078;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 CLNPITVAFI 305
 |||||
 Db 310 CLNPITVAFI 319

RESULT 26
 R70123
 ID R70123 standard; Protein; 1060 AA.
 AC R70123:
 DT 14-FEB-1996 (first entry)
 DE IL8-R type 1-GBP 130 fusion protein.
 KW Hybrid peptide; malaria parasite; Plasmodium falciparum; fusion protein;
 KW red blood cell; cytokine receptor; glycophorin binding peptide 130;
 KW GBP 130; GBPH: glycophorin binding peptide homologue; glycophorin A.
 OS Chimeric Homo sapiens.
 OS Chimeric Plasmodium falciparum.
 PN W09506737-A.
 PD 09-MAR-1995.
 PF 01-SEP-1994; G01900.
 PR 03-SEP-1993; GB-018350.
 PR 23-AUG-1994; GB-017021.
 PA (PREN/) PRENDERGAST K F.
 PI Prendergast KF;
 DR WPI: 95-115452/15.
 PT New hybrid peptide(s) for binding cytokine(s) - comprising a
 PT malaria parasite peptide capable of binding a red blood cell and
 PT a receptor peptide.
 PS Example A: Page 79-80; 93pp; English.
 CC Hybrid peptides for binding cytokines, comprising a malaria parasite
 CC (Plasmodium falciparum) peptide (capable of binding to a red blood
 CC cell (RBC)) and a receptor peptide are claimed. R70103-25 are examples
 CC of these hybrid peptides. R70123 is a fusion of interleukin 8 receptor
 CC type 1 and glycophorin binding protein (GBP) 130. The use of cytokine
 CC receptors not normally found on RBCs means that the cytokine can bind
 CC harmlessly to the RBC without deleterious effect. The RBC protects the
 CC hybrid peptides from excretion from the kidney, and due to steric
 CC hindrance prevents the cytokines binding to a receptor in another cell.
 CC GBP 130 or GBPH (GBP homologue) are the prefid. malaria parasite peptides
 CC used, others include EBA 175 (175 kda erythrocyte binding antigen).
 CC PMMSA (pre major merozoite surface antigen) and the Duffy binding
 CC receptor molecule (eg. exhibited by Plasmodium vivax). These peptides
 CC bind to pref. glycophorin A, B and C, stalo glycoproteins, found on
 CC the surface of RBCs. The hybrid peptides are thus used to lower
 CC the levels of free cytokines in the circulation to reduce pathological
 CC damage.
 SO Sequence 1060 AA:

Query Match 2.8%; Score 10; DB 1; Length 1060;
 Best Local Similarity 100.0%; Pred. No. 0.2;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 Db 134 DRYLAIVHAT 143

RESULT 27
 R70124
 ID R70124 standard; Protein; 1064 AA.
 AC R70124:
 DT 14-FEB-1996 (first entry)
 DE IL8-R type 2-GBP 130 fusion protein.
 KW Hybrid peptide; malaria parasite; Plasmodium falciparum; fusion protein;
 KW red blood cell; cytokine receptor; glycophorin binding peptide 130;
 KW GBP 130; GBPH: glycophorin binding peptide homologue; glycophorin A.
 OS Chimeric Homo sapiens.
 OS Chimeric Plasmodium falciparum.
 PN W09506737-A.
 PD 09-MAR-1995.
 PF 01-SEP-1994; G01900.
 PR 03-SEP-1993; GB-018350.
 PR 23-AUG-1994; GB-017021.
 PA (PREN/) PRENDERGAST K F.
 PI Prendergast KF;
 DR WPI: 95-115452/15.
 PT New hybrid peptide(s) for binding cytokine(s) - comprising a
 PT malaria parasite peptide capable of binding a red blood cell and
 PT a receptor peptide.
 PS Example A: Page 80-82; 93pp; English.
 CC Hybrid peptides for binding cytokines, comprising a malaria parasite
 CC (Plasmodium falciparum) peptide (capable of binding to a red blood
 CC cell (RBC)) and a receptor peptide are claimed. R70103-25 are examples
 CC of these hybrid peptides. R70124 is a fusion of interleukin 8 receptor
 CC type 2 and glycophorin binding protein (GBP) 130. The use of cytokine
 CC receptors not normally found on RBCs means that the cytokine can bind
 CC harmlessly to the RBC without deleterious effect. The RBC protects the
 CC hybrid peptides from excretion from the kidney, and due to steric
 CC hindrance prevents the cytokines binding to a receptor in another cell.
 CC GBP 130 or GBPH (GBP homologue) are the prefid. malaria parasite peptides
 CC used, others include EBA 175 (175 kda erythrocyte binding antigen).
 CC PMMSA (pre major merozoite surface antigen) and the Duffy binding
 CC receptor molecule (eg. exhibited by Plasmodium vivax). These peptides
 CC bind to pref. glycophorin A, B and C, stalo glycoproteins, found on
 CC the surface of RBCs. The hybrid peptides are thus used to lower
 CC the levels of free cytokines in the circulation to reduce pathological
 CC damage.
 SO Sequence 1064 AA:

Query Match 2.8%; Score 10; DB 1; Length 1064;
 Best Local Similarity 100.0%; Pred. No. 0.2;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 Db 137 DRYLAIVHAT 146

RESULT 28
 R80756
 ID R80756 standard; Protein; 350 AA.
 AC R80756:
 DT 26-MAR-1996 (first entry)
 DE Interleukin 8 receptor A partial sequence.
 KW Interleukin; IL-8; inflammation; psoriasis; dermatitis;
 KW rheumatoid arthritis; inflammatory bowel disease;
 KW chronic lung inflammation; treatment; antibody;
 KW affinity purification; detection.
 OS Homo sapiens.
 PN U55440021-A.

PD 08-AUG-1995: 677211.
 PF 29-MAR-1991: US-677211.
 PR 29-MAR-1991: US-677211.
 PR 25-FEB-1994: US-202056.
 PA (CHUN/) CHUNTHARAPAI A.
 PA (HEBE/) HERBERT C.
 PA (KIM/) KIM K J.
 PA (LEEJ/) LEE J.
 PI Chuntharapai A, Herbert C, Kim KJ, Lee J;
 DR WPI: 95-283151/37.
 DR N-PSDB: Q99006.
 PT New antibodies against interleukin 8 type B receptor - used to treat or prevent inflammation, also for detecting receptor expression and purification.
 PS Example 2: Columns 41-44: 62pp: English.
 CC Antibodies directed against the interleukin-8 receptor B can be used to treat or prevent inflammation e.g. psoriasis, dermatitis, rheumatoid arthritis and particularly inflammatory bowel disease and chronic lung inflammation. When immobilised, these antibodies may be used to detect interleukin-8 receptor B expression in cells and tissues and for affinity purification of interleukin-8 receptor B from cells. A sequence encoding amino acids 23-314 of this fragment of the high affinity interleukin-8 receptor, was used to probe lambda gt10 cDNA libraries for the human interleukin-8 type B receptor.
 SQ Sequence 350 AA:

Query Match 2.8%; Score 10; DB 1; Length 350;
 Best Local Similarity 100.0%; Pred. No. 0.075;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DRYLAIVHAT 142
 |||||
 DB 134 DRYLAIVHAT 143

RESULT 29
 R80758
 ID R80758 standard; Protein; 360 AA.
 AC R80758:
 DI 26-MAR-1996 (first entry)
 DE Interleukin 8 receptor B.
 KW Interleukin: IL-8; inflammation; psoriasis; dermatitis;
 KW rheumatoid arthritis; inflammatory bowel disease;
 KW chronic lung inflammation; treatment; antibody;
 KW affinity purification; detection.
 OS Homo sapiens.
 PN US5440021-A.
 PD 08-AUG-1995.
 PF 29-MAR-1991: 677211.
 PR 29-MAR-1991: US-677211.
 PR 25-FEB-1994: US-202056.
 PA (CHUN/) CHUNTHARAPAI A.
 PA (HEBE/) HERBERT C.
 PA (KIM/) KIM K J.
 PA (LEEJ/) LEE J.
 PI Chuntharapai A, Herbert C, Kim KJ, Lee J;
 DR WPI: 95-283151/37.
 DR N-PSDB: Q99008.
 PT New antibodies against interleukin 8 type B receptor - used to treat or prevent inflammation, also for detecting receptor expression and purification.
 PS Disclosure: Columns 53-56: 62pp: English.
 CC Antibodies directed against the interleukin-8 receptor B can be used to treat or prevent inflammation e.g. psoriasis, dermatitis, rheumatoid arthritis and particularly inflammatory bowel disease and chronic lung inflammation. When immobilised, these antibodies may be used to detect interleukin-8 receptor B expression in cells and tissues and for affinity purification of interleukin-8 receptor B from cells.
 SQ Sequence 360 AA:

Query Match 2.8%; Score 10; DB 1; Length 360;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DRYLAIVHAT 142
 |||||
 DB 143 DRYLAIVHAT 152

RESULT 30
 R80953
 ID R80953 standard; Protein; 360 AA.
 AC R80953:
 DI 24-APR-1996 (first entry)
 DE Recombinant high affinity interleukin-8 receptor subtype B.
 KW IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder; anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;
 KW systemic necrotising vasculitis; psoriasis; asthma; allergy; ARDS;
 KW adult respiratory distress syndrome; neutrophil detection.
 OS Homo sapiens.
 PN W09525126-A1.
 PD 21-SEP-1995.
 PF 09-MAR-1995: U03032.
 PR 15-MAR-1994: US-210250.
 PR 02-MAY-1994: US-237937.
 PA (REPK) REPLIGEN CORP.
 PA (UYBO-) UNIV BOSTON.
 PI Greenfield EA, Larosa GJ, Navarro J, Thomas KM;
 PI Wilt DT
 DR WPI: 95-336945/43.
 DR N-PSDB: Q99952.
 PT Monoclonal antibody against recombinant IL-8 receptor polypeptide - useful for treating inflammatory disorders, for detecting neutrophil(s) and for isolating IL-8 receptor from liq.mixt.
 PS Claim 5: Fig 4A-B: 74pp: English.
 CC Monoclonal antibodies were raised against recombinant interleukin-8 (IL-8) receptor subtypes A and B from both human and rabbit sources (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8rA) is a high affinity receptor and the B subtype receptor (IL-8rB) is a low affinity receptor. The monoclonal antibody (mAb) pref. binds to the IL-8 binding domain thus blocking its activation. The mAbs are useful for treating inflammatory disorders (see key words) and for detecting the presence of neutrophils in a biological sample. The mAbs are also useful in the isolation of IL-8 receptors from a mixture.
 SQ Sequence 360 AA:

Query Match 2.8%; Score 10; DB 1; Length 360;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DRYLAIVHAT 142
 |||||
 DB 143 DRYLAIVHAT 152

RESULT 31
 R80950
 ID R80950 standard; Protein; 355 AA.
 AC R80950:
 DI 24-APR-1996 (first entry)
 DE Recombinant high affinity interleukin-8 receptor subtype A.
 KW IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder; anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;
 KW systemic necrotising vasculitis; psoriasis; asthma; allergy; ARDS;
 KW adult respiratory distress syndrome; neutrophil detection.
 OS Oryzotagus cuniculus.
 PN W09525126-A1.
 PD 21-SEP-1995.
 PF 09-MAR-1995: U03032.
 PR 15-MAR-1994: US-210250.
 PR 02-MAY-1994: US-237937.
 PA (REPK) REPLIGEN CORP.

PA (UYBO-) UNIV BOSTON.
 PI Greenfield EA, Larosa GJ, Navarro J, Thomas KM;
 PI Wilt DT;
 DR WPI: 95-336945/43.
 DR N-PSDB: Q99949.
 PT Monoclonal antibody against recombinant IL-8 receptor polypeptide -
 useful for treating inflammatory disorders, for detecting
 PI neutrophil(s) and for isolating IL-8 receptor from liq.mixt.
 PS Claim 2: Fig 1A-B: 74pp: English.
 CC Monoclonal antibodies were raised against recombinant interleukin-8
 CC (IL-8) receptor subtypes A and B from both human and rabbit sources
 CC (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8RA) is
 CC a high affinity receptor and the B subtype receptor (IL-8RB) is a
 CC low affinity receptor. The monoclonal antibody (mAb) pref. binds to
 CC the IL-8 binding domain thus blocking its activation. The mAbs are
 CC useful for treating inflammatory disorders (see key words) and for
 CC detecting the presence of neutrophils in a biological sample. The
 CC mAbs are also useful in the isolation of IL-8 receptors from a mixture.
 SQ Sequence 355 AA;

Query Match 2.8%; Score 10; DB 1; Length 355;
 Best Local Similarity 100.0%; Pred. No. 0.076;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 133 DRYLAIVHAT 142
 Db 139 DRYLAIVHAT 148

RESULT 32
 ID R80951
 AC R80951: standard; Protein: 350 AA.
 DE 24-APR-1996 (first entry)
 DE Recombinant high affinity interleukin-8 receptor subtype A.
 DE IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder;
 DE anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;
 DE systemic necrotizing vasculitis; psoriasis; asthma; allergy; ARDS;
 DE adult respiratory distress syndrome; neutrophil detection.
 OS Homo sapiens.
 OS W09525126-A1.
 PD 21-SEP-1995.
 PR 09-MAR-1995: U03032.
 PR 15-MAR-1994: US-210250.
 PR 02-MAY-1994: US-237937.
 PA (REPK) REPLIGEN CORP.
 PA (UYBO-) UNIV BOSTON.
 PI Greenfield EA, Larosa GJ, Navarro J, Thomas KM;
 PI Wilt DT;
 DR WPI: 95-336945/43.
 DR N-PSDB: Q99950.
 PT Monoclonal antibody against recombinant IL-8 receptor polypeptide -
 useful for treating inflammatory disorders, for detecting
 PI neutrophil(s) and for isolating IL-8 receptor from liq.mixt.
 PS Disclosure: Fig 2A-B: 74pp: English.
 CC Monoclonal antibodies were raised against recombinant interleukin-8
 CC (IL-8) receptor subtypes A and B from both human and rabbit sources
 CC (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8RA) is
 CC a high affinity receptor and the B subtype receptor (IL-8RB) is a
 CC low affinity receptor. The monoclonal antibody (mAb) pref. binds to
 CC the IL-8 binding domain thus blocking its activation. The mAbs are
 CC useful for treating inflammatory disorders (see key words) and for
 CC detecting the presence of neutrophils in a biological sample. The
 CC mAbs are also useful in the isolation of IL-8 receptors from a mixture.
 SQ Sequence 350 AA;

Query Match 2.8%; Score 10; DB 1; Length 350;
 Best Local Similarity 100.0%; Pred. No. 0.075;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 133 DRYLAIVHAT 142

Db 134 DRYLAIVHAT 143

RESULT 33
 ID R80952
 AC R80952: standard; Protein: 358 AA.
 DE 24-APR-1996 (first entry)
 DE Recombinant high affinity interleukin-8 receptor subtype B.
 DE IL-8A; IL-8B; receptor; monoclonal antibody; inflammatory disorder;
 DE anaphylaxis; systemic lupus erythematosus; rheumatoid arthritis;
 DE systemic necrotizing vasculitis; psoriasis; asthma; allergy; ARDS;
 DE adult respiratory distress syndrome; neutrophil detection.
 OS Oryzolaqus cuniculus.
 OS W09525126-A1.
 PD 21-SEP-1995.
 PR 09-MAR-1995: U03032.
 PR 15-MAR-1994: US-210250.
 PR 02-MAY-1994: US-237937.
 PA (REPK) REPLIGEN CORP.
 PA (UYBO-) UNIV BOSTON.
 PI Greenfield EA, Larosa GJ, Navarro J, Thomas KM;
 PI Wilt DT;
 DR WPI: 95-336945/43.
 DR N-PSDB: Q99951.
 PT Monoclonal antibody against recombinant IL-8 receptor polypeptide -
 useful for treating inflammatory disorders, for detecting
 PI neutrophil(s) and for isolating IL-8 receptor from liq.mixt.
 PS Claim 6: Fig 3A-B: 74pp: English.
 CC Monoclonal antibodies were raised against recombinant interleukin-8
 CC (IL-8) receptor subtypes A and B from both human and rabbit sources
 CC (R80950-53 encoded by Q99949-52). The A subtype receptor (IL-8RA) is
 CC a high affinity receptor and the B subtype receptor (IL-8RB) is a
 CC low affinity receptor. The monoclonal antibody (mAb) pref. binds to
 CC the IL-8 binding domain thus blocking its activation. The mAbs are
 CC useful for treating inflammatory disorders (see key words) and for
 CC detecting the presence of neutrophils in a biological sample. The
 CC mAbs are also useful in the isolation of IL-8 receptors from a mixture.
 SQ Sequence 358 AA;

Query Match 2.8%; Score 10; DB 1; Length 358;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 133 DRYLAIVHAT 142
 Db 141 DRYLAIVHAT 150

RESULT 34
 ID W25217
 AC W25217: standard; peptide: 367 AA.
 DE 12-JAN-1998 (first entry)
 DE Rat orphanin FQ receptor encoded by cDNA clone LC132.
 DE Orphanin FQ receptor; binding; locomotor disease; diagnosis;
 DE treatment; opioid inhibitor; opiate induced hypothermia; drug design;
 DE morphine induced analgesia; methadone specific opioid receptor; ss.
 OS Rattus sp.
 OS Key
 FT Misc_difference 21 Location/Qualifiers
 FT Misc_difference 26 /note= "potential N-linked glycosylation site"
 FT Misc_difference 36 /note= "potential N-linked glycosylation site"
 FT Misc_difference 45..75 /note= "potential N-linked glycosylation site"
 FT Domain 81..106 /note= "transmembrane domain I"
 FT Domain 122..143 /note= "transmembrane domain II"
 FT Domain

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 1999, 02:59:26 ; Search time 11.06 Seconds

(without alignments)
1275.140 Million cell updates/sec

Title: US-09-104-063-4

Sequence: 1 MEGSITYSNDNTEEMSGSD.....KRGHSSVSTESSESSFHSS 352

Scoring table: OLIGO

Searched: 122810 segs, 40065486 residues

Database : PIR_60:*

1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	352	100.0	352	2	A45747	neuropeptide Y/pep
2	114	32.4	352	2	G00048	fusin (LESTRA) - c
3	74	21.0	353	2	S28787	neuropeptide Y/pep
4	10	2.8	350	2	A39445	interleukin-8 rece
5	10	2.8	360	2	A53611	interleukin-8 rece
6	10	2.8	370	2	S43087	orphan opioid rece
7	10	2.8	358	2	A53752	interleukin-8 rece
8	10	2.8	355	2	JQ1231	interleukin-8 rece
9	10	2.8	359	2	A48921	interleukin-8 rece
10	10	2.8	367	2	I49022	K3 opiate receptor
11	10	2.8	367	2	JC2421	opioid receptor ho
12	10	2.8	367	2	I56520	G protein-coupled
13	10	2.8	356	2	S42096	interleukin-8 rece

ALIGNMENTS

RESULT 1
A45747
neuropeptide Y/peptide YY receptor Y3 - human
N:Alternate names: fusin; HM99; leukocyte-derived seven-transmembrane receptor LESTR; re
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 17-Mar-1999
C:Accession: A45747; A53103; I53006; I59444; I69203; S33761
R:Feierspiel, B.; Melhado, I.G.; Duncan, A.M.V.; Delaney, A.; Schappert, K.; Clark-Lewis
Genomics 16, 707-712, 1993
A:Title: Molecular cloning of the cDNA and chromosomal localization of the gene for a pu
A:Reference number: A45747; MUID:93315154
A:Accession: A45747
A:Molecule type: mRNA
A:Residues: 1-352 <FE0>
A:Cross-references: GB:M99293; NID:9292516; PID:9292517
R:Loetscher, M.; Geisler, T.; O'Reilly, T.; Zechlen, R.; Baggiolini, M.; Moser, B.
J. Biol. Chem. 269, 232-237, 1994
A:Title: Cloning of a human seven-transmembrane domain receptor, LESTR, that is highly e
A:Reference number: A53103; MUID:94103215
A:Accession: A53103
A:Molecule type: mRNA

A:Residues: 1-352 <LOE>
A:Cross-references: EMBL:X71635; NID:9297099; PID:9297100
R:Herzog, H.; Hort, T.O.; Shine, J.; Seibler, L.A.
DNA Cell Biol. 12, 465-471, 1993
A:Title: Molecular cloning, characterization, and localization of the human homolog t
A:Reference number: I53006; MUID:933196529
A:Accession: I53006

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-352 <HER>

A:Cross-references: GB:L06797; NID:9414929; PID:9414928
R:Jazin, E.E.; Yoo, H.; Blomqvist, A.G.; Yee, F.; Weng, G.; Walker, M.W.; Salton, J.;
Regul. Pept. 47, 247-258, 1993

A:Title: A proposed bovine neuropeptide Y (NPY) receptor cDNA clone, or its human hom

A:Reference number: I59444; MUID:94052833

A:Accession: I59444

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-352 <RE2>

A:Cross-references: GB:L01639; NID:9189313; PID:9189314
R:Nomura, H.; Nielsen, B.W.; Matsushima, K.
Int. Immunol. 5, 1239-1249, 1993

A:Title: Molecular cloning of cDNAs encoding a LD78 receptor and putative leukocyte c

A:Reference number: I54751; MUID:94092629

A:Accession: I69203

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-352 <RES>

A:Cross-references: GB:ID10924; NID:9219868; PID:9219869

A:Gene: GDB:NPY3R; NPY3

A:Cross-references: GDB:230002; OMIM:162643

A:Map position: 2921-2921

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match	100.0%	Score 352;	DB 2;	Length 352;
Best local similarity	100.0%	Pred. No. 0;		
Matches 352;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MEGSITYSNDNTEEMSGSDYDMSKPCFREANFNKIFLPTIYSIFLGTIVGNGLVI	60	
DB	1	MEGSITYSNDNTEEMSGSDYDMSKPCFREANFNKIFLPTIYSIFLGTIVGNGLVI	60	
QY	61	LVWGYCKLRSMYDKYRLSLVADLLFVTLPFWADAVANNYFGNFKAVHYITVNL	120	
DB	61	LVWGYCKLRSMYDKYRLSLVADLLFVTLPFWADAVANNYFGNFKAVHYITVNL	120	
QY	121	YSSVLLIAFISLDRLAIYHATNSQRPRLAEKVYVGVWIPALLITIPDFIFANVSEA	180	
DB	121	YSSVLLIAFISLDRLAIYHATNSQRPRLAEKVYVGVWIPALLITIPDFIFANVSEA	180	
QY	181	DDRYTCDREYFNDLVWVYQFOHIMWGLLPBGIVLISCCIIISKISHSGQKRAKLT	240	
DB	181	DDRYTCDREYFNDLVWVYQFOHIMWGLLPBGIVLISCCIIISKISHSGQKRAKLT	240	
QY	241	TYLLIAFAFACMLPYPIGISISFILLEIKOGCEPENVHWMISTELAFHCCLNPI	300	
DB	241	TYLLIAFAFACMLPYPIGISISFILLEIKOGCEPENVHWMISTELAFHCCLNPI	300	
QY	301	LYAFLGAKKRTSAOHALTSVSRGSSLKITSKGRGSHSVSTESSESSFHSS	352	
DB	301	LYAFLGAKKRTSAOHALTSVSRGSSLKITSKGRGSHSVSTESSESSFHSS	352	

A:Reference number: H00048
 A:Accession: G00048
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-352 <TAT>
 A:Cross-references: GB:D86579; NID:g1468948; PID:g1468949

Query Match 32.4%; Score 114; DB 2; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1.8e-107;
 Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 177 VSEADRRICDFRYPNDLVVVFQFIHWGLILPGIVILSCYCIISKLSHGKGRK 236
 |||||||
 Db 177 VSEADRRICDFRYPNDLVVVFQFIHWGLILPGIVILSCYCIISKLSHGKGRK 236
 |||||||
 Oy 237 ALKTVIILAFACMLPYIGISIDSFLLEIKOGCEFEVTHKWSITLAL 290
 |||||||
 Db 237 ALKTVIILAFACMLPYIGISIDSFLLEIKOGCEFEVTHKWSITLAL 290
 |||||||

RESULT 3
 S28787
 neuropeptide Y/peptide YY receptor Y3 - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 06-Dec-1996
 C:Accession: S28787
 R:Rimland, J.; Xin, W.; Sweetnam, P.; Saijoh, K.; Nestler, E.J.; Duman, R.S.
 Mol. Pharmacol. 40, 869-875, 1991
 A:Title: Sequence and expression of a neuropeptide Y receptor cDNA.
 A:Reference number: S28787; MUID:92100053
 A:Accession: S28787
 A:Molecule type: mRNA
 A:Residues: 1-353 <RIM>
 A:Cross-references: EMBL:M86739
 C:Keywords: appetite; G protein-coupled receptor; transmembrane protein

Query Match 21.0%; Score 74; DB 2; Length 353;
 Best Local Similarity 100.0%; Pred. No. 5.6e-67;
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 279 TVHKKISTEALAFPHCCNPLIYAFILGAKFKTSQAHLTSYRSGSLILSKGRGHS 338
 |||||||
 Db 280 TVHKKISTEALAFPHCCNPLIYAFILGAKFKTSQAHLTSYRSGSLILSKGRGHS 339
 |||||||
 Oy 339 SVSTESSESSFHSS 352
 |||||||
 Db 340 SVSTESSESSFHSS 353
 |||||||

RESULT 4
 A39445
 Interleukin-8 receptor type A - human
 N:Alternate names: Interleukin-8 receptor, high-affinity
 C:Species: Homo sapiens (man)
 C:Date: 22-Jan-1993 #sequence_revision 12-Apr-1996 #text_change 24-Sep-1998
 C:Accession: I37449; I38710; I38711; A39445
 R:Moilleau, C.; Muscatelli, F.; Mattei, M.G.; Vassart, G.; Parmentier, M.
 Genomics 16, 248-251, 1993
 A:Title: The high-affinity Interleukin 8 receptor gene (IL8RA) maps to the 2q33-q36 reg1
 A:Reference number: I37449; MUID:93252387
 A:Accession: I37449
 A:Molecule type: DNA
 A:Residues: 1-350 <RES>
 A:Cross-references: EMBL:X65858; NID:g312046; PID:g312047
 R:Ahuja, S.K.; Shetty, A.; Tiffany, H.L.; Murphy, P.M.
 J. Biol. Chem. 269, 26381-26389, 1994
 A:Title: Comparison of the genomic organization and promoter function for human interleukin-8 receptor type A
 A:Reference number: I37898; MUID:95014476
 A:Accession: I38710
 A:Molecule type: DNA
 A:Residues: 1-350 <REZ>

A:Cross-references: EMBL:U11870; NID:g511804; PID:g511805
 A:Accession: I38711
 A:Molecule type: mRNA
 A:Residues: 1-16 <RE3>
 A:Cross-references: EMBL:U11871; NID:g511806; PID:g733002
 R:Holmes, W.E.; Lee, J.; Kuang, W.J.; Rice, G.C.; Wood, W.I.
 Science 253, 1278-1280, 1991
 A:Title: Structure and functional expression of a human interleukin-8 receptor.
 A:Reference number: A39445; MUID:91368199
 A:Accession: A39445
 A:Molecule type: mRNA
 A:Residues: 1-275, 'T', 277-350 <HOL>
 A:Cross-references: GB:M68932; NID:g186369; PID:g186370
 C:Genetics:
 A:Gene: GDB:IL8RA
 A:Cross-references: GDB:I35039; OMIM:146929
 A:Map position: 2q35-2q35
 C:Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 2.8%; Score 10; DB 2; Length 350;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 133 DRYLAIVHAT 142
 |||||||
 Db 134 DRYLAIVHAT 143
 |||||||

RESULT 5
 A53611
 Interleukin-8 receptor type B - human
 C:Species: Homo sapiens (man)
 C:Date: 07-Oct-1994 #sequence_revision 12-Apr-1996 #text_change 29-Aug-1997
 C:Accession: I37898; I38712; A53611; A39446
 R:Ahuja, S.K.; Shetty, A.; Tiffany, H.L.; Murphy, P.M.
 J. Biol. Chem. 269, 26381-26389, 1994
 A:Title: Comparison of the genomic organization and promoter function for human interleukin-8 receptor type B
 A:Reference number: I37898; MUID:95014476
 A:Accession: I37898
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-360 <RES>
 A:Cross-references: EMBL:U11869; NID:g511801; PID:g511803
 A:Accession: I38712
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-15 <RE2>
 A:Cross-references: EMBL:U11872; NID:g511808; PID:g511809; EMBL:U11873; NID:g511810;
 6; PID:g511817; EMBL:U11877; NID:g511818; PID:g511819; EMBL:U11878; NID:g511820; PID:
 R:Sprengr, H.; Lloyd, A.R.; Lautens, L.L.; Bonner, T.L.; Kelyin, D.J.
 J. Biol. Chem. 269, 11065-11072, 1994
 A:Title: Structure, genomic organization, and expression of the human interleukin-8 r

A:Reference number: A53611; MUID:94209273
 A:Accession: A53611
 A:Molecule type: DNA
 A:Status: preliminary
 A:Residues: 6-360 <SPR>
 A:Cross-references: GB:M99412; GB:U19593
 R:Murphy, P.M.; Tiffany, H.L.
 Science 253, 1280-1283, 1991
 A:Title: Cloning of complementary DNA encoding a functional human interleukin-8 recep
 A:Reference number: A39446; MUID:91368200
 A:Accession: A39446
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 6-360 <MUR>
 A:Cross-references: GB:M73969
 C:Comment: This receptor, unlike IL8RA, binds several peptides besides interleukin-8,
 C:Genetics:
 A:Gene: GDB:IL8RB; IL8RA
 A:Cross-references: GDB:I27868; OMIM:146928
 A:Map position: 2q35-2q35

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 2.8%; Score 10; DB 2; Length 360;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 133 DRYLAIVHAT 142
|||||
Db 143 DRYLAIVHAT 152

RESULT 6

S43087
Orphan opioid receptor ORL1 - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 10-Sep-1997
C:Accession: S43087; G01892
R:Moilleau, C.; Parmentier, M.; Mailleux, P.; Butour, J.L.; Moisand, C.; Chalon, P.; Ca
FEBS Lett. 341, 33-38, 1994
A:Title: ORL1, a novel member of the opioid receptor family. Cloning, functional express
A:Reference number: S43087; MUID:94185768
A:Accession: S43087
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-370 <MOL>
A:Cross-references: EMBL:X77130; NID:g471316; PID:g471317
R:Lee P., H.
submitted to the EMBL Data Library, June 1995
A:Reference number: H00703
A:Accession: G01892
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-370 <LEF>
A:Cross-references: EMBL:U30185; NID:g1144296; PID:g1144297
C:Genetics:
A:Gene: GDB:OPRL1
A:Cross-references: GDB:345029
A:Map position: Bq11.2-Bq11.2
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 2.8%; Score 10; DB 2; Length 370;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 296 CLNPITVAF 305
|||||
Db 313 CLNPITVAF 322

RESULT 7

A53752
Interleukin-8 receptor (clone 5B1a) - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Sep-1997
C:Accession: A53752
R:Prado, G.N.; Thomas, K.M.; Suzuki, H.; LaRosa, G.J.; Wilkinson, N.; Folco, E.; Navaric
J. Biol. Chem. 269, 12391-12394, 1994
A:Title: Molecular characterization of a novel rabbit interleukin-8 receptor isotype.
A:Reference number: A53752; MUID:94230294
A:Accession: A53752
A:Molecule type: mRNA
A:Status: preliminary
A:Residues: 1-358 <PRA>
A:Cross-references: GB:L24445; NID:g437661; PID:g437662
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 2.8%; Score 10; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 133 DRYLAIVHAT 142
|||||
Db 141 DRYLAIVHAT 150

RESULT 8

J01231
Interleukin-8 receptor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Mar-1998
C:Accession: J01231; A46483
R:Beckmann, M.P.; Munger, W.E.; Kozlosky, C.; Vandembos, T.; Price, V.; Lyman, S.; Ge
Biochem. Biophys. Res. Commun. 179, 784-789, 1991
A:Title: Molecular characterization of the Interleukin-8 receptor.
A:Reference number: J01231; MUID:91378994
A:Accession: J01231
A:Molecule type: DNA
A:Residues: 1-355 <BEC>
A:Cross-references: GB:M74240; NID:g165438; PID:g165439
R:Lee, J.; Kuang, W.J.; Rice, G.C.; Wood, W.I.
J. Immunol. 148, 1261-1264, 1992
A:Title: Characterization of complementary DNA clones encoding the rabbit IL-8 recept
A:Reference number: A46483; MUID:92148149
A:Accession: A46483
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-355 <LEF>
A:Cross-references: GB:M82873; NID:g165440; PID:g165441
A:Experimental source: neutrophils
A:Note: sequence extracted from NCBI backbone (NCBI:81526, NCBI:81530)
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 2.8%; Score 10; DB 2; Length 355;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 133 DRYLAIVHAT 142
|||||
Db 139 DRYLAIVHAT 148

RESULT 9

A48921
Interleukin-8 receptor type B - mouse
N:Alternate names: G-protein coupled receptor Gpcr16
C:Species: Mus musculus (house mouse)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 17-Mar-1999
C:Accession: A48921; A53677; I49348; I5421; H48909; I53774
R:Cerretti, D.P.; Nelson, N.; Kozlosky, C.J.; Morrissey, P.J.; Copeland, N.G.; Gilber
Genomics 18, 410-413, 1993
A:Title: The murine homologue of the human interleukin-8 receptor type B maps near th
A:Reference number: A48921; MUID:94117014
A:Accession: A48921
A:Molecule type: DNA
A:Residues: 1-359 <CER>
A:Cross-references: GB:L23637; NID:g435093; PID:g435094
R:Suzuki, H.; Prado, G.N.; Wilkinson, N.; Navarro, J.
J. Biol. Chem. 269, 18263-18266, 1994
A:Title: The N terminus of interleukin-8 (IL-8) receptor confers high affinity bindin
A:Reference number: A53677; MUID:94308043
A:Accession: A53677
A:Molecule type: DNA
A:Status: nucleic acid sequence not shown: not compared with conceptual translation
A:Residues: 1-359 <SUZ>
A:Cross-references: GB:L26549
A:Note: sequence extracted from NCBI backbone (NCBI:149812)
R:Lee, J.; Caclalano, G.; Camerato, T.; Toy, K.; Moore, M.W.; Wood, W.I.
J. Immunol. 155, 2158-2164, 1995
A:Title: Chemokine binding and activities mediated by the mouse IL-8 receptor.
A:Reference number: I49348; MUID:95363183
A:Accession: I49348
A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA
 A:Residues: 1-359 <RES>
 A:Cross-references: EMBL:U01207; NID:9550174; PID:9550175
 R:Bozic, C.R.; Gerard, N.P.; von Döbeln-Guldenband, C.; Kolakowski, L.F.
 J. Biol. Chem. 269, 29355-29358, 1994
 A:Title: The murine interleukin 8 type B receptor homologue and its ligands. Expression
 A:Reference number: 155421; MUID:95050766
 A:Accession: 155421
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-359 <RE2>
 A:Cross-references: GB:J13339; NID:9293665; PID:9293666
 R:Wilke, T.M.; Chen, Y.; Gillebert, D.J.; Moore, K.J.; Yu, L.; Simon, M.I.; Copeland, N.G.
 Genomics 18, 175-184, 1993
 A:Title: Identification, chromosomal location, and genome organization of mammalian G-protein coupled receptor
 A:Reference number: A48909; MUID:94116980
 A:Accession: H48909
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 145-258 <MIL>
 A:Cross-references: GB:J20337; NID:9438800; PID:9438801
 R:Harada, A.; Kuno, K.; Nomura, H.; Mukaida, N.; Murakami, S.; Matsushima, K.
 Gene 142, 297-300, 1994
 A:Title: Cloning of a cDNA encoding a mouse homolog of the interleukin-8 receptor.
 A:Reference number: 153774; MUID:94252584
 A:Accession: 153774
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-359 <RE3>
 A:Cross-references: GB:J17630; NID:9493671; PID:9493672
 A:Gene: 118-b
 A:Genetics:
 A:Introns: #status absent
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F:49-74/Domain: transmembrane #status predicted <TM1>
 F:84-106/Domain: transmembrane #status predicted <TM2>
 F:120-141/Domain: transmembrane #status predicted <TM3>
 F:163-182/Domain: transmembrane #status predicted <TM4>
 F:213-234/Domain: transmembrane #status predicted <TM5>
 F:251-271/Domain: transmembrane #status predicted <TM6>
 F:308-328/Domain: transmembrane #status predicted <TM7>

Query Match 2.8%; Score 10; DB 2; Length 359;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 DRYLATVHAT 142
 |||||
 Db 142 DRYLATVHAT 151

RESULT 10
 K3 opiate receptor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 02-Jul-1996
 C:Accession: J49022
 R:Pan, Y.X.; Cheng, J.; Xu, J.; Rossi, G.; Jacobson, E.; Ryan-Moro, J.; Brooks, A.I.; De
 Mol. Pharmacol. 47, 1180-1188, 1995
 A:Title: Cloning and functional characterization through antisense mapping of a kappa 3-
 A:Reference number: J49022; MUID:95327076
 A:Accession: J49022
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-367 <RES>
 A:Cross-references: EMBL:U09421; NID:9551484; PID:9551485

Query Match 2.8%; Score 10; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 CLNPITLYAFL 305
 |||||
 Db 310 CLNPITLYAFL 319

RESULT 11
 JC2421
 Opioid receptor homolog, MOR-C - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 21-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 13-Mar-1997
 C:Accession: JC2421; J49122
 R:Nishi, M.; Takeshima, H.; Mori, M.; Nakagawara, K.; Takeuchi, T.
 Biochem. Biophys. Res. Commun. 205, 1353-1357, 1994
 A:Title: Structure and chromosomal mapping of genes for the mouse kappa-opioid recept
 A:Reference number: JC2434; MUID:95100967
 A:Accession: JC2421
 A:Molecule type: mRNA
 A:Residues: 1-367 <NIS>
 A:Cross-references: DDBJ:D31663
 R:Halford, W.P.; Gebhardt, B.M.; Carr, D.J.
 J. Neuroimmunol. 59, 91-101, 1995
 A:Title: Functional role and sequence analysis of a lymphocyte orphan opioid receptor
 A:Reference number: J49122; MUID:95318231
 A:Accession: J49122
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-357 <RES>
 A:Cross-references: EMBL:U14165; NID:9540092; PID:9540093
 C:Genetics:
 A:Gene: MOR-C; OOR
 A:Map position: 2H2-4
 A:Introns: 75/2; 194/1
 C:Keywords: receptor

Query Match 2.8%; Score 10; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 CLNPITLYAFL 305
 |||||
 Db 310 CLNPITLYAFL 319

RESULT 12
 G protein-coupled receptor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 29-Jan-1999
 C:Accession: I56520; S46238; S43655; S45701; I52654
 R:Lachowicz, J.E.; Shen, Y.; Monsma, F.J.
 J. Neurochem. 64, 34-40, 1995
 A:Title: Molecular cloning of a novel G protein-coupled receptor related to the opiat
 A:Reference number: I56520; MUID:95096849
 A:Accession: I56520
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-367 <RES>
 A:Cross-references: EMBL:U07871; NID:9606802; PID:9606803
 R:Chen, Y.; Fan, Y.; Liu, J.; Mestek, A.; Tian, M.; Kozak, C.A.; Yu, L.
 FEBS Lett. 347, 279-283, 1994
 A:Title: Molecular cloning, tissue distribution and chromosomal localization of a nov
 A:Reference number: S46238; MUID:94307400
 A:Accession: S46238
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-367 <CHE>
 A:Cross-references: GB:J28144; NID:9496219; PID:9496220
 R:Fukuda, K.; Kato, S.; Mori, K.; Nishi, M.; Takeshima, H.; Iwabe, N.; Miyata, T.; Ho
 FEBS Lett. 343, 42-46, 1994
 A:Title: cDNA cloning and regional distribution of a novel member of the opioid recep
 A:Reference number: S43655; MUID:94215703
 A:Accession: S43655

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-367 <FEU>
 A:Cross-references: GB:D16438; NID:g456168; PID:81004424; PID:g533355
 R:Bunzow, J.R.; Saez, C.; Mottard, M.; Bouvier, C.; Williams, J.T.; Low, M.; Grandy, D.K.
 FEBS Lett. 347, 284-288, 1994
 A>Title: Molecular cloning and tissue distribution of a putative member of the rat opioi
 A:Reference number: S46235; MUID:94307401
 A:Accession: S46235
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-347, 'A', 349-367 <BUN>
 A:Cross-references: EMBL:U01913; NID:g487964; PID:g487965
 A>Note: the authors translated the codon GCG for residue 348 as Ser
 R:Wang, J.B.; Johnson, P.S.; Imel, Y.; Persico, A.M.; Ozenberger, B.A.; Eppler, C.M.; UH
 FEBS Lett. 348, 75-79, 1994
 A>Title: cDNA cloning of an orphan opiate receptor gene family member and its splice var
 A:Reference number: S45701; MUID:94298959
 A:Accession: S45701
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-331, 'LL', 334-367 <WAN>
 R:Wick, M.J.; Minnerath, S.R.; Lin, X.; Elide, R.; Law, P.Y.; Loh, H.H.
 Brain Res. Mol. Brain Res. 27, 37-44, 1994
 A>Title: Isolation of a novel cDNA encoding a putative membrane receptor with high homol
 A:Reference number: I52654; MUID:95182817
 A:Accession: I52654
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-367 <RE2>
 A:Cross-references: GB:I29419; NID:g510718; PID:g510719
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 2.8%; Score 10; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.036;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 CUNPILYAF 305
 |||||
 DB 310 CUNPILYAF 319

RESULT 13
 S42096
 Interleukin-8 receptor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Feb-1997
 C:Accession: S42096
 R:GOBL, A.E.; Wang, S.; Zhou, Y.; Oeberg, K.
 submitted to the EMBL Data Library, February 1994
 A:Description: Molecular cloning of the rat IL8 receptor.
 A:Reference number: S42096
 A:Accession: S42096
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-356 <GOB>
 A:Cross-references: EMBL:X77797
 C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 2.8%; Score 10; DB 2; Length 356;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 DB 142 DRYLAIVHAT 151

Thu Aug 12 09:17:08 1999

us-09-104-063-4.rpt

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 1999, 05:59:13 ; Search time 10.2 Seconds
(without alignments)
975,533 Million cell updates/sec

Title: US-09-104-063-4
Sequence: 352
1 MEGASITSDMYTEEMSGD.....KRGHSSVSTSESSSFHSS 352

Scoring table: OLIGO

Searched: 77977 seqs, 28268293 residues

Database : SwissProt_37.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	352	100.0	352	1	CCR4_HUMAN
2	213	60.5	352	1	CCR4_PAPAN
3	168	47.7	352	1	CCR4_CERIO
4	138	39.2	352	1	CCR4_MACMU
5	114	32.4	352	1	CCR4_MACFA
6	74	21.0	353	1	CCR4_BOVIN
7	74	21.0	353	1	CCR4_FELCA
8	60	17.0	359	1	CCR4_MOUSE
9	60	17.0	349	1	CCR4_RAT
10	55	15.6	192	1	CCR4_SHEEP
11	10	2.8	350	1	IL8A_GORCO
12	10	2.8	350	1	IL8A_HUMAN
13	10	2.8	350	1	IL8A_PANTR
14	10	2.8	355	1	IL8A_RABIT
15	10	2.8	349	1	IL8A_RAT
16	10	2.8	360	1	IL8B_BOVIN
17	10	2.8	353	1	IL8B_GORCO
18	10	2.8	360	1	IL8B_HUMAN
19	10	2.8	353	1	IL8B_MACMU
20	10	2.8	359	1	IL8B_MOUSE
21	10	2.8	353	1	IL8B_PANTR
22	10	2.8	358	1	IL8B_RABIT
23	10	2.8	359	1	IL8B_RAT
24	10	2.8	370	1	OPRX_CAVPO
25	10	2.8	370	1	OPRX_HUMAN
26	10	2.8	367	1	OPRX_MOUSE
27	10	2.8	370	1	OPRX_PIG
28	10	2.8	367	1	OPRX_RAT

ALIGNMENTS

RESULT 1
ID CCR4_HUMAN STANDARD: PRT: 352 AA.
AC P30991: P56438: ..
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LEUKOCYTE-DERIVED
DE SEVEN TRANSMEMBRANE DOMAIN RECEPTOR) (LESTR) (LCR1) (FB22) (NPYRL)

DE (HM89).
GN CXCR4.
OS HOMO SAPIENS (HUMAN), AND PAN TROGLODITES (CHIMPANZEE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-HUMAN: TISSUE=LUNG;
RX MEDLINE: 93319629.
RA HERZOG H., HORT Y.J., SHINE J., SELBIE L.A.;
RT "Molecular cloning, characterization, and localization of the human
RT homolog to the reported bovine NPY Y3 receptor: lack of NPY binding
RT and activation."
RN DNA CELL BIOL. 12:465-471(1993).
[2]
RP SEQUENCE FROM N.A.
RC SPECIES-HUMAN: TISSUE=BRAIN;
RX MEDLINE: 94052833.
RA JAZIN E.E., YOO H., BLOMOVIST G., YEE F., WENG G., WALKER M.W.,
RA SALON J., LARHAMMAR D., WAHLESTEDT C.R.;
RT "A proposed bovine neuropeptide Y (NPY) receptor cDNA clone, or its
RT human homologue, confers neither NPY binding sites nor NPY
RT responsiveness on transfected cells."
RN REGUL. PEPT. 47:247-258(1993).
[3]
RP SEQUENCE FROM N.A.
RC SPECIES-HUMAN: TISSUE=SPLEEN;
RX MEDLINE: 93315164.
RA FEDERSPIETEL B., DELANEY A.D., CLARK-LEWIS I., JIRIK F., DUNCAN A.M.,
RA SCHAPPERT K.T., MELHADO I.;
RT "Molecular cloning of the cDNA and chromosomal localization of the
RT gene for a putative seven-transmembrane segment (7-TMS) receptor
RT isolated from human spleen."
RN GENOMICS 16:707-712(1993).
[4]
RP SEQUENCE FROM N.A.
RC SPECIES-HUMAN: TISSUE=LEUKOCYTE;
RX MEDLINE: 94103215.
RA LOETSCHER M., GEISER T., O'REILLY T., ZWAHLEN R., BAGGIOLINI M.,
RA MOSER B.;
RT "Cloning of a human seven-transmembrane domain receptor, LESTR, that
RT is highly expressed in leukocytes."
RN J. BIOL. CHEM. 269:232-237(1994).
[5]
RP SEQUENCE FROM N.A.
RC SPECIES-HUMAN: TISSUE=MONOCYTES;
RX MEDLINE: 94092629.
RA NOMURA H., NIELSEN B.W., MATSUSHIMA K.;
RT "Molecular cloning of cDNAs encoding a Ld78 receptor and putative
RT leukocyte chemotactic peptide receptors."
RN INT. IMMUNOL. 5:1239-1249(1993).
[6]
RP SEQUENCE FROM N.A.
RC SPECIES-HUMAN;
RX WEGNER S.A., EHRENBERG P.K., CHANG G., DAYHOFF D.E., MICHAEL N.L.;
RN SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
[7]
RP SEQUENCE FROM N.A.
RC SPECIES=P. TROGLODITES;
RX MEDLINE: 98090115.
RA PRETER J. L., ZERBIB A., GIRARD M., GUILLET J.-G., BUTOR C.;
RT "Chimpanzee CXCR4 and CCR5 act as coreceptors for HIV type 1."
RN AIDS RES. HUM. RETROVIRUSES 13:1583-1587(1997).
[8]
RP FUNCTION.
RX MEDLINE: 96351077.
RA BLEUL C.C., FARZAN M., CHOE H., PAROLIN C., CLARK-LEWIS I.,
RA SODROSKI J., SPRINGER T.A.;
RT "The lymphocyte chemotactant SDF-1 is a ligand for LESTR/Fusin and
RT blocks HIV-1 entry."
RN NATURE 382:829-833(1996).
[9]
RP FUNCTION.

FT	DISURFID	11	186	POTENTIAL.	BY SIMILARITY.
SO	SEQUENCE	352 AA:	39745 MW:	BDE02133 CRC32:	
Query Match		100.0%;	Score 352;	DB 1;	Length 352:
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches	352;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
QY	1	MEGISITSDNYTEEMSGSDYDSKMEPCFEENANFKIFLPITYSIETLTGIVGNLVI	60		
Db	1	MEGISITSDNYTEEMSGSDYDSKMEPCFEENANFKIFLPITYSIETLTGIVGNLVI	60		
QY	61	LVMGYOKKLRSMTDKRYRLHLVSADLFEVILLPFMAVADVAMNYFGNLFCAVHYITVNL	120		
Db	61	LVMGYOKKLRSMTDKRYRLHLVSADLFEVILLPFMAVADVAMNYFGNLFCAVHYITVNL	120		
QY	121	YSSVILLAFISLDRIYLAIVAHATNSQRPKRLLAEVYVVGWVPIALLITIDPFIFANSEA	180		
Db	121	YSSVILLAFISLDRIYLAIVAHATNSQRPKRLLAEVYVVGWVPIALLITIDPFIFANSEA	180		
QY	181	DDRYICDAEFYNDLVVVFQFQILMVGILIPGIYILSCYCIILSKLSHGOKRKALKT	240		
Db	181	DDRYICDAEFYNDLVVVFQFQILMVGILIPGIYILSCYCIILSKLSHGOKRKALKT	240		
QY	241	TVILLIAEFACWLPYPIIGISIDSFITLLEIKQCFEFNYAKVISTEALAFHCCINPI	300		
Db	241	TVILLIAEFACWLPYPIIGISIDSFITLLEIKQCFEFNYAKVISTEALAFHCCINPI	300		
QY	301	LYAFIGAFKTSQAQHALTVSRGSSKLITSGKRGHSHVSTEESSSFHSS	352		
Db	301	LYAFIGAFKTSQAQHALTVSRGSSKLITSGKRGHSHVSTEESSSFHSS	352		
RESULT	2				
ID	CCRA.PAPAN	STANDARD;	PRT:	352 AA.	
AC	P56491:				
DT	15-JUL-1998 (REL. 36, CREATED)				
DT	15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)				
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)				
DE	C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (CXCR-4) (SDF-1 RECEPTOR)				
DE	(STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN).				
GN	CXCR4.				
OS	PAP10 ANUBIS (OLIVE BABOON).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;				
OC	PRIMATES; CATARRHINI; CERCOPITHECIDAE; CERCOPITHECINAE; PAP10.				
RP	[1]				
RA	SEQUENCE FROM N.A.				
RA	BENTON P.A., TIMANUS D.K., SHEARER M.H., LEE D.R., KENNEDY R.C.:				
RL	SUBMITTED (NOV-1997) TO EMBL/GENBANK/DDJ DATA BANKS.				
CC	-1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A				
CC	SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.				
CC	-1- SUBCELLULAR LOCATION: INTERAL MEMBRANE PROTEIN.				
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation				
CC	the European Bioinformatics Institute. There are no restrictions on ways				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC					
DR	EMBL: AF031089; G2625094; .				
DR	GCRDB: GCR_2512; .				
DR	POSTITE: PS00237; G-PROTEIN RECEPTOR; 1.				
KW	G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRAN; GLYCOPROTEIN.				
FT	DOMAIN	1	39	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	40	63	1 (POTENTIAL).	
FT	DOMAIN	64	79	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM	80	99	2 (POTENTIAL).	
FT	DOMAIN	100	110	EXTRACELLULAR (POTENTIAL).	


```

FT TRANSMEM 111 132 3 (POTENTIAL).
FT DOMAIN 133 134 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 155 175 4 (POTENTIAL).
FT DOMAIN 176 200 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 201 220 5 (POTENTIAL).
FT DOMAIN 221 240 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 241 261 6 (POTENTIAL).
FT DOMAIN 262 285 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 286 305 7 (POTENTIAL).
FT TRANSMEM 306 352 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 11 11 POTENTIAL.
FT DISULFID 109 186 BY SIMILARITY.
SQ SEQUENCE 352 AA: 39751 MM: 273DB8EE CRC32:

Query Match
Best Local Similarity 60.5%; Score 213; DB 1; Length 352:
Matches 313; Conservative 0; Mismatches 1; Indels 0; Gaps 0:

QY 39 ILEPTIVSIFLGIYVNGVILVMGYQKRLRSMTDKYRLHLSVADLLFVITLPEFAVDA 98
DB 39 ILEPTIVSIFLGIYVNGVILVMGYQKRLRSMTDKYRLHLSVADLLFVITLPEFAVDA 98
QY 99 VANWYFGNFKRAVHYIYVNLKXSVLLAFISLDRLALVHATNSORPKLLAEKVVVY 158
DB 99 VANWYFGNFKRAVHYIYVNLKXSVLLAFISLDRLALVHATNSORPKLLAEKVVVY 158
QY 159 GWAIPALLITPDIFFANVSEADRYICDRFPNDLWVWVVOFOHIMVGLILPGIIVLSC 218
DB 159 GWAIPALLITPDIFFANVSEADRYICDRFPNDLWVWVVOFOHIMVGLILPGIIVLSC 218
QY 219 YCIIILKLSKSGHOKRAKKTIVILLIAFFACMLPYIGISIDSFTLLIINOGCEFFEN 278
DB 219 YCIIILKLSKSGHOKRAKKTIVILLIAFFACMLPYIGISIDSFTLLIINOGCEFFEN 278
QY 279 TVHKWISITELALFFHCCLNPILYAFLGAKFKTSAQHALLTSVSRGSSSLKILSKGRGHS 338
DB 279 TVHKWISITELALFFHCCLNPILYAFLGAKFKTSAQHALLTSVSRGSSSLKILSKGRGHS 338
QY 339 SVSTESESSSFHSS 352
DB 339 SVSTESESSSFHSS 352

RESULT 3
CCR4_CERTO
ID CCR4_CERTO STANDARD: PRT: 352 AA.
AC 062747:
DT 15-DEC-1998 (REL. 37, CREATED)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LESTR).
GN CXCR4.
OS CEROCCEBUS TOROUATUS AIYS (RED-CROWNED MANGABEY) (SOOTY MANGABEY).
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
OC PRIMATES: CATARRHINI: CERCOPITHECIDAE: CERCOPITHECINAE: CERCOCEBUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 98321155.
RA CHEN Z., GETTIE A., HO D.D., MARX P.A.:
"Primary stvsm isolates use the CCR5 coreceptor from sooty mangabeys
naturally infected in west Africa: a comparison of coreceptor usage
of primary stvsm. HIV-2, and stvmac."
VIRIDIC 246:113-124(1998).
RL
CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF051906; G3135304; -
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT DOMAIN 1 39
FT TRANSMEM 40 63 1 (POTENTIAL).
FT DOMAIN 64 79 2 (POTENTIAL).
FT TRANSMEM 80 99 3 (POTENTIAL).
FT DOMAIN 100 110 4 (POTENTIAL).
FT TRANSMEM 111 132 5 (POTENTIAL).
FT DOMAIN 133 154 6 (POTENTIAL).
FT TRANSMEM 155 175 7 (POTENTIAL).
FT DOMAIN 176 200 8 (POTENTIAL).
FT TRANSMEM 201 220 9 (POTENTIAL).
FT DOMAIN 221 240 10 (POTENTIAL).
FT TRANSMEM 241 261 11 (POTENTIAL).
FT DOMAIN 262 285 12 (POTENTIAL).
FT TRANSMEM 286 305 13 (POTENTIAL).
FT DOMAIN 306 352 14 (POTENTIAL).
FT CARBOHYD 11 11 POTENTIAL.
FT DISULFID 109 186 BY SIMILARITY.
SQ SEQUENCE 352 AA: 39648 MM: 8B464DAD CRC32:

Query Match
Best Local Similarity 47.7%; Score 168; DB 1; Length 352:
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 185 ICDRFYNDLWVWVVOFOHIMVGLILPGIIVLSCYCIILKLSKSGHOKRAKKTIVIL 244
DB 185 ICDRFYNDLWVWVVOFOHIMVGLILPGIIVLSCYCIILKLSKSGHOKRAKKTIVIL 244
QY 245 ILAFFACMLPYIGISIDSFTLLIINOGCEFFENTYHKWISITELALFFHCCLNPILYAF 304
DB 245 ILAFFACMLPYIGISIDSFTLLIINOGCEFFENTYHKWISITELALFFHCCLNPILYAF 304
QY 305 LGAKFKTSAQHALLTSVSRGSSSLKILSKGRGHSVSTESESSSFHSS 352
DB 305 LGAKFKTSAQHALLTSVSRGSSSLKILSKGRGHSVSTESESSSFHSS 352

RESULT 4
CCR4_MACMU
ID CCR4_MACMU STANDARD: PRT: 352 AA.
AC P79394; 002745; 046428:
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LESTR).
GN CXCR4.
OS MACACA MULATTA (RHESUS MACAQUE).
OC EUKARYOTA: METAZOA: CHORDATA: VERTEBRATA: MAMMALIA: EUTHERIA:
OC PRIMATES: CATARRHINI: CERCOPITHECIDAE: CERCOPITHECINAE: MACACA.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-INDIAN MACAQUE;
RX MEDLINE: 97213934.
RA CHEN Z., ZHOU P., HO D.D., LANDAU N.R., MARX P.A.:
"Genetically divergent strains of simian immunodeficiency virus use
CCR5 as a coreceptor for entry."
J. VIROL. 71:2705-2714(1997).
RL [2]
CC SEQUENCE FROM N.A.
CC MEDLINE: 9726687.
CC EDINGER A.L., AMEDEE A., MILLER K., DORANZ B.J., ENDRES M.,
CC SHARON M., SAMSON M., LU Z.-H., CLEMENTS J.E., MORRHEE-CORB M.,
CC PEIPER S.C., PARMENTIER M., BRODER C.C., DOMS R.W.:
CC "Differential utilization of CCR5 by macrophage and T cell tropic

```

RT simian immunodeficiency virus strains".
RL PROC. NATL. ACAD. SCI. U.S.A. 94:4005-4010(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CHINESE.
RA MEDLINE: 98252393.
RX PRETER J.-L., GUILLET J.-G., BUTOR C.:
RT "New widespread CXCR4 allele in rhesus macaques does not predict
RT subspecies or clinical evolution."
RL AIDS RES. HUM. RETROVIRUSES 14:639-641(1998).
CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U73740; G1899056; -
DR EMBL: U93311; G1934671; -
DR EMBL: AF001928; G2911294; -
DR GCRDB: GCR_1297; -
DR GCRDB: GCR_1640; -
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm_1; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT DOMAIN 1 39
FT TRANSMEM 40 63
FT DOMAIN 64 79
FT TRANSMEM 80 99
FT DOMAIN 100 110
FT TRANSMEM 111 132
FT DOMAIN 133 154
FT TRANSMEM 155 175
FT DOMAIN 176 200
FT TRANSMEM 201 220
FT DOMAIN 221 240
FT TRANSMEM 241 261
FT DOMAIN 262 285
FT TRANSMEM 286 305
FT DOMAIN 306 352
FT CARBOHYD 11
FT DISULFID 109 186
FT CONFLICT 67 67
FT CONFLICT 214 214
FT CONFLICT 348 348
SQ SEQUENCE 352 AA; 39739 MW; 3C33A4BB CRC32;
Query Match 39.2%; Score 138; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.4e-131;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ID CCR4_MACEFA STANDARD; PRT: 352 AA.
AC 028474;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LESTR).
GN CXCR4.
OS MACACA FASCICULARIS (CRAB EATING MACAQUE) (CYNOMOLGUS MONKEY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; CERCOPITHECIDAE; CERCOPITHECINAE; MACACA.
RN [1]
RP SEQUENCE FROM N.A.
RA TATSUMI M., TAKAHASHI H.;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDSI DATA BANKS.
CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D86579; G1468949; -
DR GCRDB: GCR_1143; -
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm_1; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT DOMAIN 1 39
FT TRANSMEM 40 63
FT DOMAIN 64 79
FT TRANSMEM 80 99
FT DOMAIN 100 110
FT TRANSMEM 111 132
FT DOMAIN 133 154
FT TRANSMEM 155 175
FT DOMAIN 176 200
FT TRANSMEM 201 220
FT DOMAIN 221 240
FT TRANSMEM 241 261
FT DOMAIN 262 285
FT TRANSMEM 286 305
FT DOMAIN 306 352
FT CARBOHYD 11
FT DISULFID 109 186
SQ SEQUENCE 352 AA; 39753 MW; 7EDA93BA CRC32;
Query Match 32.4%; Score 114; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 3.7e-107;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 177 VSEADRYTCDFRPYDMLVWVYFQOHVWGLILGIVLSCYCIILSKLSKSHKORR 236
DB 177 VSEADRYTCDFRPYDMLVWVYFQOHVWGLILGIVLSCYCIILSKLSKSHKORR 236
QY 237 ALKTVIILAFACWLPYIGISIDSFILLEIKOGCFEFTVHKWISTEAL 290
DB 237 ALKTVIILAFACWLPYIGISIDSFILLEIKOGCFEFTVHKWISTEAL 290
RESULT 6
CCR4_BOVIN STANDARD; PRT: 353 AA.
AC P25930;
DT 01-MAY-1992 (REL. 22, CREATED)
DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

Query Match 21.0%; Score 74; DB 1; Length 353;
 Best Local Similarity 100.0%; Pred. No. 7,7e-67;
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 279 TVHKMISTEALAFHCCNPLTYAFLEKFKTSQHALTSVRSGLKLSKGRGHS 338
 |||
 Db 280 TVHKMISTEALAFHCCNPLTYAFLEKFKTSQHALTSVRSGLKLSKGRGHS 339
 |||
 Y 339 SVTSESSSESSFHSS 352
 |||
 Db 340 SVTSESSSESSFHSS 353

RESULT 8
 CCR4_MOUSE STANDARD; PRT: 359 AA.
 ID CCR4_MOUSE P70658; P70346: 009062; 009059; P70233;
 AC 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (SDF-1 RECEPTOR)
 DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LEUKOCYTE-DERIVED
 SEVEN TRANSMEMBRANE DOMAIN RECEPTOR) (LESTR).
 GN CXCR4 OR LESTR OR CMKAR4 OR SDF1R.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 RN RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MOS.
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=THYMUS;
 RA MOEPFS B., FRODL R., KESSLER H., GIERSCHEK P.;
 RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SV;
 RA HESEN M., BERMAN M.A., GERARD C., DORF M.E.;
 RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA MEDLINE: 97113334.
 RX HESEN M., BERMAN M.A., BENSON J.D., GERARD C., DORF M.E.;
 RT "Cloning of the mouse fusin gene, homologue to a human HIV-1
 co-factor";
 RT J. IMMUNOL. 157:5455-5460(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BONE MARROW;
 RX MEDLINE: 97121456.
 RA NAGASAWA T., NAKAUMA T., TACHIBANA K., ITZASA H., BLEUL C.C.,
 RA YOSHIE O., MATSUSHIMA K., YOSHIDA N., SPRINGER T.A., KISHIMOTO T.;
 RT "Molecular cloning and characterization of a murine pre-B-cell
 growth-stimulating factor/stromal cell-derived factor 1 receptor, a
 murine homolog of the human immunodeficiency virus 1 entry coreceptor
 fusin";
 RT PROC. NATL. ACADE. SCI. U.S.A. 93:14726-14729(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=THYMUS;
 RA SUZUKI G., NAKATA Y., UZAWA A., SHIRASAWA T., SAITO T., MITA K.;
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SV; TISSUE=THYMUS;
 RA SCHUBEL A., BURGSTAHLER R., LIPP M.;
 RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS;
 CC -1- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
 SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X99581; E281127; -;
 DR EMBL: X99582; E281126; -;
 DR EMBL: U59760; G1527135; -;
 DR EMBL: U65580; G1731651; -;
 DR EMBL: D87747; G1772446; -;
 DR EMBL: AB000803; G1816446; -;
 DR EMBL: 280111; E265602; -;
 DR EMBL: 280112; E181578; -;
 DR GCRDB; GCR_1138; -;
 DR GCRDB; GCR_1387; -;
 DR GCRDB; GCR_1646; -;
 DR GCRDB; GCR_1730; -;
 DR GCRDB; GCR_2592; -;
 DR MGD; MGI:109563; CMKAR4.
 DR PROSITE; PS00237; G_PROTEIN_RECEPTOR; 1.
 DR PFM; PF00001; 7tm_1; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
 FT DOMAIN 1 41
 FT TRANSMEM 42 65
 FT DOMAIN 66 81
 FT TRANSMEM 82 101
 FT DOMAIN 102 112
 FT TRANSMEM 113 134
 FT DOMAIN 135 156
 FT TRANSMEM 157 177
 FT DOMAIN 178 207
 FT TRANSMEM 208 227
 FT DOMAIN 228 247
 FT TRANSMEM 248 268
 FT DOMAIN 269 292
 FT TRANSMEM 293 312
 FT DOMAIN 313 359
 FT DISULFID 111 193
 FT CARBOHYD 13 13
 FT CONFLICT 6 7
 FT CONFLICT 216 216
 FT SEQUENCE 359 AA; 40426 MW; 1037B4D3 CRC32;

Query Match 17.0%; Score 60; DB 1; Length 359;
 Best Local Similarity 100.0%; Pred. No. 1e-52;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 197 VVFOFHIMVGLIPGIVILSCYCIITISKLSHSGHOKRKALKTIVILLAFACWLPY 256
 |||
 Db 204 VVFOFHIMVGLIPGIVILSCYCIITISKLSHSGHOKRKALKTIVILLAFACWLPY 263
 |||

RESULT 9
 CCR4_RAT STANDARD; PRT: 349 AA.
 ID CCR4_RAT 008565;
 AC 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (SDF-1 RECEPTOR)
 DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LEUKOCYTE-DERIVED
 SEVEN TRANSMEMBRANE DOMAIN RECEPTOR) (LESTR).
 GN CXCR4 OR CMKAR4.
 OS RATTUS NOREUGIUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WISTAR; TISSUE=SPLEEN;

```

RA HARRISON J.K., SALAFRANCA M.N.:
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U90610; G1906613; -.
CC GCRDB: GCR_1401; -.
DR GCRDB: GCR_1401; -.
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm_1; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT DOMAIN 1 36
FT TRANSMEM 37 60
FT DOMAIN 61 76
FT TRANSMEM 77 96
FT DOMAIN 97 107
FT TRANSMEM 108 129
FT DOMAIN 130 151
FT TRANSMEM 152 172
FT DOMAIN 173 197
FT TRANSMEM 198 217
FT DOMAIN 218 237
FT TRANSMEM 238 258
FT DOMAIN 259 282
FT TRANSMEM 283 302
FT DOMAIN 303 349
FT DISULFID 106 183
FT CARBOHYD 8
SQ SEQUENCE 349 AA; 39334 MW; CBA6532 CRC32;

Query Match 17.0%; Score 60; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 9.8e-53;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 VVEQFQHWGLIPGIVLSCYCIITISKLSHSGHOKRAKATVYLLIAFAACMPY 256
DB 194 VVFQFQHWGLIPGIVLSCYCIITISKLSHSGHOKRAKATVYLLIAFAACMPY 253

RESULT 10
CCR4_SHEEP STANDARD: PRT: 192 AA.
AC 028553;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE C-X-C CHEMOKINE RECEPTOR TYPE 4 (CXCR-4) (CXCR-4) (SDF-1 RECEPTOR)
DE (STROMAL CELL-DERIVED FACTOR 1 RECEPTOR) (FUSIN) (LEUCOCYTE-DERIVED
DE SEVEN TRANSMEMBRANE DOMAIN RECEPTOR) (LESTR) (FRAGMENT).
GN CXCR4 OR LESTR.
OS OVIS ARIES (SHEEP).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVIDAE; CAPRINAE; OVIS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-HYPOTHALAMUS;
RA DYER C.J., MATTERI R.L., KETSLER D.H.;
RT "Development of an ovine Y3 cDNA and expression of the Y3 receptor
RT mRNA in the ovine hypothalamus and pituitary.";
RL ABSTR. SOC. NEUROSCI. 21:1890-1890(1995).
CC -!- FUNCTION: RECEPTOR FOR THE C-X-C CHEMOKINE SDF-1. TRANSDUCES A
CC SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS LEVEL.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

```

```

CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- CAUTION: WAS ORIGINALLY THOUGHT TO BE A RECEPTOR FOR NEUROPEPTIDE
CC Y, TYPE 3 (NPY3-R).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U38942; G1061412; -.
CC GCRDB: GCR_1381; -.
DR GCRDB: GCR_1381; -.
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm_1; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN.
FT NON_TER 1 1
FT DOMAIN <1 29
FT TRANSMEM 30 53
FT DOMAIN 54 69
FT TRANSMEM 70 89
FT DOMAIN 90 100
FT TRANSMEM 101 122
FT DOMAIN 123 144
FT TRANSMEM 145 165
FT DOMAIN 166 190
FT TRANSMEM 191 >192
FT DISULFID 99 176
FT NON_TER 192 192
SQ SEQUENCE 192 AA; 22178 MW; 6B3DEBD5 CRC32;

Query Match 15.6%; Score 55; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 6.3e-48;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 107 FLCKAVHYTVNLVSSVLIATISIDRYLAIVHANSQPRKLLAEKVVYGVW 161
DB 97 FLCKAVHYTVNLVSSVLIATISIDRYLAIVHANSQPRKLLAEKVVYGVW 151

RESULT 11
IL8A_GORGO STANDARD: PRT: 350 AA.
AC P55919; P55921;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR A (IL-8R A) (IL-8 RECEPTOR TYPE
DE 1) (CXCR-1) (CDW128).
GN IL8RA OR CXCR1.
OS GORILLA GORILLA GORILLA (LOWLAND GORILLA), AND
OS PONGO PYGMAEUS (ORANGUTAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; GORILLA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96175151.
RA ALVAREZ V., COTO E., SETIEN F., GONZALEZ S., GONZALEZ-ROCES S.,
RA LOPEZ-LARREA C.;
RT "Characterization of interleukin-8 receptors in non-human primates.";
RL IMMUNOGENETICS 43:261-267(1996).
CC -!- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
CC AND TO MGSA (GRO) WITH A LOW AFFINITY.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


DR EMBL: M74240; G155439; -
 DR EMBL: M58021; G155443; -
 DR EMBL: M82873; G155441; -
 DR PIR: A23659; A23669;
 DR PIR: A46483; A46483;
 DR PIR: J01231; J01231;
 DR GCRDB: GCR-0107; -
 DR GCRDB: GCR-0108; -
 DR GCRDB: GCR-0298; -
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1
 DR G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 40
 FT TRANSMEM 41 67
 FT DOMAIN 68 73
 FT TRANSMEM 74 92
 FT DOMAIN 93 114
 FT TRANSMEM 115 138
 FT TRANSMEM 139 159
 FT TRANSMEM 160 184
 FT TRANSMEM 185 204
 FT TRANSMEM 205 232
 FT TRANSMEM 233 247
 FT TRANSMEM 248 270
 FT TRANSMEM 271 290
 FT TRANSMEM 291 313
 FT TRANSMEM 314 355
 FT CARBOHYD 7 21
 FT CARBOHYD 21 192
 FT DISULFID 90 111
 FT CONFLICT 146 147
 FT CONFLICT 204 204
 FT CONFLICT 287 288
 FT SEQUENCE 355 AA; 40622 MW; 2EB3947D CRC32;

Query Match 2.8%; Score 10; DB 1; Length 355;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||||||
 Db 139 DRYLAIVHAT 148

RESULT 15
 IL8A_RAT STANDARD: PRT: 349 AA.
 ID IL8A_RAT
 AC P70612;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR A (IL-8R A) (CXCR-1).
 GN IL8RA OR CXCR1.
 OS RATIUS NORVEGICUS (RAT).
 OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIROGNATHI; MORIDAE; MORINAE; RATIUS.
 CC [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=MTSTAR; TISSUE=LUNG;
 RC MEDLINE: 97115810.
 RA DUNSTAN C.-A.N., SALAFRANCA M.N., ADHIKARI S., XIA Y., FENG L.,
 RA HARRISON J.K.;
 RT "Identification of two rat genes orthologous to the human
 interleukin-8 receptors."
 RT J. Biol. Chem. 271:32770-32776(1996).
 RL -i- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND

CC MESSENGER SYSTEM (BY SIMILARITY).
 CC -i- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -i- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U71089; G1589931; -
 DR GCRDB: GCR-1404; -
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1.
 DR HSSP: P34986; 1DDD.
 DR G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 44
 FT TRANSMEM 45 71
 FT TRANSMEM 72 80
 FT TRANSMEM 81 101
 FT TRANSMEM 102 116
 FT TRANSMEM 117 138
 FT TRANSMEM 139 159
 FT TRANSMEM 160 179
 FT TRANSMEM 180 204
 FT TRANSMEM 205 225
 FT TRANSMEM 226 247
 FT TRANSMEM 248 269
 FT TRANSMEM 270 290
 FT TRANSMEM 291 313
 FT TRANSMEM 314 349
 FT CARBOHYD 22 22
 FT DISULFID 115 192
 FT SEQUENCE 349 AA; 39944 MW; A8415DD5 CRC32;

Query Match 2.8%; Score 10; DB 1; Length 349;
 Best Local Similarity 100.0%; Pred. No. 0.024;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||||||
 Db 139 DRYLAIVHAT 148

RESULT 16
 IL8B_BOVIN STANDARD: PRT: 360 AA.
 ID IL8B_BOVIN
 AC Q28003;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2).
 GN IL8RB OR CXCR2.
 OS BOS TAURUS (BOVINE).
 OS EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 CC ARTIODACTYLA; RUMINANTIA; PECORA; BOVIDAE; BOVIDAE; BOS.
 CC [1]
 RN SEQUENCE FROM N.A.
 RP LI Y., FENG J., TEMPLETON J.W.;
 RL SUBMITTED (JAN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -i- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 CC AND TO GRO/MGSA AND MAP-2 ALSO WITH A HIGH AFFINITY (BY
 CC SIMILARITY).
 CC -i- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -i- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.


```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U19947; G1145878;
DR GCRDB: GCR_1224;
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm_1; 1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW CHEMOTAXIS.
FT DOMAIN 1 48 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 49 75 1 (POTENTIAL).
FT DOMAIN 76 84 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 85 105 2 (POTENTIAL).
FT DOMAIN 106 120 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 121 142 3 (POTENTIAL).
FT DOMAIN 143 163 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 164 183 4 (POTENTIAL).
FT DOMAIN 184 208 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 209 231 5 (POTENTIAL).
FT DOMAIN 232 251 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 252 273 6 (POTENTIAL).
FT DOMAIN 274 294 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 295 315 7 (POTENTIAL).
FT DOMAIN 316 360 CYTOPLASMIC (POTENTIAL).
FT DISULFID 119 196 BY SIMILARITY.
FT CARBOHYD 10 10 POTENTIAL.
FT SEQUENCE 360 AA; 40625 MW; 281DD168 CRC32;
SO
Query Match 2.8%; Score 10; DB 1; Length 360;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 133 DRYLAIVHAT 142
Db 143 DRYLAIVHAT 152
RESULT 17
IL8B_GORGO
ID IL8B_GORGO STANDARD: PRT: 353 AA.
AC Q28422;
DT 01-NOV-1997 (REL. 35, CREATED)
DI 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (FRAGMENT).
GN IL8RB OR CXCR2.
OS GORILLA GORILLA (LOWLAND GORILLA).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; GORILLA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96175151.
RA ALVAREZ V., COTO E., SETIEN F., GONZALEZ S., GONZALEZ-ROCES S.,
RA LOPEZ-LARREA C.;
RT "Characterization of interleukin-8 receptors in non-human primates.";
RL IMMUNOGENETICS 43:261-267(1995).
CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X91114; E198175;
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
DR PFAM: PF00001; 7tm_1; 1.
DR HSSP: P34996; 1DD1.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW CHEMOTAXIS.
FT NON_TER 1 1
FT DOMAIN <1 45 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 46 72 1 (POTENTIAL).
FT DOMAIN 73 81 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 82 102 2 (POTENTIAL).
FT DOMAIN 103 117 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 118 139 3 (POTENTIAL).
FT DOMAIN 140 160 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 161 180 4 (POTENTIAL).
FT DOMAIN 181 205 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 206 228 5 (POTENTIAL).
FT DOMAIN 229 248 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 249 270 6 (POTENTIAL).
FT DOMAIN 271 291 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 292 312 7 (POTENTIAL).
FT DOMAIN 313 >353 CYTOPLASMIC (POTENTIAL).
FT DISULFID 116 193 BY SIMILARITY.
FT CARBOHYD 19 19 POTENTIAL.
FT NON_TER 353 353
FT SEQUENCE 353 AA; 39919 MW; 4AF43313 CRC32;
SO
Query Match 2.8%; Score 10; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 133 DRYLAIVHAT 142
Db 140 DRYLAIVHAT 149

```

```

RESULT 18
IL8B_HUMAN
ID IL8B_HUMAN STANDARD: PRT: 360 AA.
AC P25025;
DT 01-MAY-1992 (REL. 22, CREATED)
DI 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DI 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (GRO/MGSA
DE RECEPTOR) (IL-8 RECEPTOR TYPE 2).
GN IL8RB OR CXCR2.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91368200.
RA MURPHY P.M., TIFFANY H.L.;
RT "Cloning of complementary DNA encoding a functional human
RT interleukin-8 receptor.";
RL SCIENCE 253:1280-1283(1991).
CC [2]
CC SEQUENCE FROM N.A., AND CHARACTERIZATION.
CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----

```

RL MOL. IMMUNOL. 30:359-367(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 94209273.
 RA SPRINGER H., LLOYD A.R., LAUTENS L.L., BONNER T.I., KELVIN D.J.;
 RT "Structure, genomic organization, and expression of the human
 RL interleukin-8 receptor B gene".
 RN J. BIOL. CHEM. 269:11065-11072(1994).
 RP [4]
 RP SEQUENCE FROM N.A.
 RX TISSUE-PLENTA;
 RC MEDLINE: 95014476.
 RA AHUJA S.K., SHETTY A., TIFFANY H.L., MURPHY P.M.;
 RT "Comparison of the genomic organization and promoter function for
 RL human interleukin-8 receptors A and B".
 RN J. BIOL. CHEM. 269:26381-26389(1994).
 RN [5]
 RP CHARACTERIZATION.
 RX MEDLINE: 92355587.
 RA LEE J., HORUK R., RICE G.C., BENNETT G.L., CAMERATO T., WOOD W.I.;
 RT "Characterization of two high affinity human Interleukin-8
 RL receptors".
 RN J. Biol. Chem. 267:16283-16287(1992).
 CC -I- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOATRACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 CC AND TO GRG/MSCA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -I- SUBCELLULAR LOCATION: INTERNAL MEMBRANE PROTEIN.
 CC -I- SIMILARITY: BELONGS TO FAMILY I OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: M73969; GI109691; -
 DR EMBL: M94582; GI186378; -
 DR EMBL: M99412; G576679; -
 DR EMBL: L19593; G559054; -
 DR EMBL: U11869; G511803; -
 DR PIR: A39446; A39446.
 DR PIR: A53611; A53611.
 DR GCRDB: GCR_0077; -
 DR GCRDB: GCR_0610; -
 DR GCRDB: GCR_1001; -
 DR GCRDB: GCR_1339; -
 DR GCRDB: GCR_1831; -
 DR MIM: 146928; -
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1.
 DR HSP: P34996; IDDD.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 48
 FT TRANSEM 49 75 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 76 84 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 85 105 2 (POTENTIAL).
 FT DOMAIN 106 120 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 121 142 3 (POTENTIAL).
 FT DOMAIN 143 163 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 164 183 4 (POTENTIAL).
 FT DOMAIN 184 208 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 209 231 5 (POTENTIAL).
 FT DOMAIN 232 251 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 252 273 6 (POTENTIAL).
 FT DOMAIN 274 294 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 295 315 7 (POTENTIAL).
 FT DOMAIN 316 360 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 316 360

```

Query Match
Best Local Similarity 100.0%: Pred. No. 0.025;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Qy 133 DRYLAVHAT 142
Db 143 DRYLAVHAT 152

RESULT 19
ID 118B_MACMU STANDARD; PRT: 353 AA.
118B_MACMU
AC 028519;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (FRAGMENT).
DS IL8RB OR CXCR2.
GN MACACA MULATTA (RHESUS MACAQUE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; CERCOPITHECIDAE; CERCOPITHECINAE; MACACA.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE: 9617511.
RA ALVAREZ V., COTO E., SETIEN F., GONZALEZ S., GONZALEZ-ROCES S.,
RA LOPEZ-LARREA C.;
RT "Characterization of interleukin-8 receptors in non-human primates.";
RL IMMUNOGENETICS 43:261-267(1996).
CC -I- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
CC AND TO GPCR/MGSA AND MAP-2 ALSO WITH A HIGH AFFINITY.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X91116; E198176; -
DR PROSITE; PS00237; G-PROTEIN_RECEPTOR; 1.
DR PFMF; PF00001; 7tm_1; 1.
DR HSSP; P34996; IDDD.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW CHEMOTAXIS.
KW
FT NON_TER 1 1
FT DOMAIN <1 45
FT TRANSSEM 46 72
FT DOMAIN 73 81
FT TRANSSEM 82 102
FT DOMAIN 103 117
FT TRANSSEM 118 139
FT DOMAIN 140 160
FT TRANSSEM 161 180
FT DOMAIN 181 205
FT TRANSSEM 206 228
FT DOMAIN 229 248
FT TRANSSEM 249 270
FT DOMAIN 271 291
FT TRANSSEM 292 312
FT DOMAIN 313 >353
FT DISULFID 116 193
FT BY SIMILARITY

```

FT CARBOHYD 19 19 POTENTIAL.
 FT NON_TER 353 353
 SO SEQUENCE 353 AA: 35947 MW: 8592FE64 CRC32:

Query Match 2.8%; Score 10; DB 1; Length 353;
 Best Local Similarity 100.0%; Pred No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 |||||
 Db 140 DRYLAIVHAT 149

RESULT 20
 ID IL8B_MOUSE STANDARD; PRT: 359 AA.
 AC P35343;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (GRO/MSA RECEPTOR).
 DE IL8RB OR CXCR2 OR CXKR2 OR GPCR16.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIOGOGNATHI; MORIDAE; MORINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 94117014.
 RA CERRETTI D.P., NELSON N., KOZLOSKY C.J., MORRISSEY P.J.,
 RA COPELAND N.G., GILBERT D.J., JENKINS N.A., DOSIK J.K., MOCK B.A.,
 RT "The murine homologue of the human interleukin-8 receptor type B maps near the lty-lsh-Bog disease resistance locus.";
 RL GENOMICS 18:410-413(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE: 95050765.
 RA BOZIC C.R., GERARD N.P., VON UEXKULL-GULDENBAND C.,
 RA KOLAKOWSKI L.F. JR., CONKLYN M.J., BRESLOW R., SHOWELL H.J.,
 RA GERARD C.,
 RT "The murine interleukin 8 type B receptor homologue and its ligands. Expression and biological characterization.";
 RL J. BIOL. CHEM. 269:29355-29358(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-129/SV;
 RX MEDLINE: 95363183.
 RA LEE J., CACALANO G., CAMERATO T., TOY K., MOORE M.W., WOOD W.I.,
 RT "Chemokine binding and activities mediated by the mouse IL-8 receptor.";
 RL J. IMMUNOL. 155:2158-2164(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 94252584.
 RA HARADA A., KUNO K., NOMURA H., MUKAIDA N., MURAKAMI S., MATSUSHIMA K.,
 RT "Cloning of a cDNA encoding a mouse homolog of the interleukin-8 receptor.";
 RL GENE 142:297-300(1994).
 RN [5]
 RP SEQUENCE OF 145-258 FROM N.A.
 RC TISSUE-TESTIS;
 RX MEDLINE: 94116980.
 RA WILKIE T.M., CHEN Y., GILBERT D.J., MOORE K.J., YU L., SIMON M.I.,
 RA COPELAND N.G., JENKINS N.A.,
 RT "Identification, chromosomal location, and genome organization of mammalian G-protein-coupled receptors.";
 RL GENOMICS 18:175-184(1993).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND

CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 CC AND TO GRO/MSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L23637; G435094; -
 DR EMBL: L13239; G293666; -
 DR EMBL: U31207; G950175; -
 DR EMBL: D17630; G493672; -
 DR EMBL: L20337; G438801; -
 DR PIR: A53677; A53677.
 DR GCRDB: GCR_0550; -
 DR GCRDB: GCR_0812; -
 DR GCRDB: GCR_0838; -
 DR GCRDB: GCR_1027; -
 DR GCRDB: GCR_1670; -
 DR MGD: MGI:105303; CXKR2.
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR: 1.
 DR PIR: PF00001; 7tm_1; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 47
 FT TRANSEM 48 74
 FT DOMAIN 75 83
 FT TRANSEM 84 104
 FT DOMAIN 105 119
 FT TRANSEM 120 141
 FT DOMAIN 142 162
 FT TRANSEM 163 182
 FT DOMAIN 183 207
 FT TRANSEM 208 230
 FT DOMAIN 231 250
 FT TRANSEM 251 272
 FT DOMAIN 273 293
 FT TRANSEM 294 314
 FT DOMAIN 315 359
 FT CARBOHYD 223 223
 FT DISULFID 118 195
 SO SEQUENCE 359 AA: 40425 MW: D501949A CRC32:
 Query Match 2.8%; Score 10; DB 1; Length 359;
 Best Local Similarity 100.0%; Pred No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 133 DRYLAIVHAT 142
 |||||
 Db 142 DRYLAIVHAT 151
 RESULT 21
 ID IL8B_PANTR STANDARD; PRT: 353 AA.
 AC Q28807;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (FRAGMENT).
 DE IL8RB OR CXCR2.
 OS PAN TROGLODYTES (CHIMPANZEE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC PRIMATES; CATARRHINI; HOMINIDAE; PAN.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 96175151.

RA ALVAREZ V., COTO E., SETIEN F., GONZALEZ S., GONZALEZ-ROCES S.,
 RA LOPEZ-LARREA C.:
 RT "Characterization of interleukin-8 receptors in non-human primates."
 RL IMMUNOGENETICS 43:261-267(1996).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X91113; E198174.
 DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1.
 DR HSSP: P34996; 1DDD.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT NON_TER 1 1
 FT DOMAIN <1 45 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 46 72 1 (POTENTIAL).
 FT DOMAIN 73 81 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 82 102 2 (POTENTIAL).
 FT DOMAIN 103 117 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 118 139 3 (POTENTIAL).
 FT DOMAIN 140 160 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 161 180 4 (POTENTIAL).
 FT DOMAIN 181 205 5 (POTENTIAL).
 FT TRANSMEM 206 228 6 (POTENTIAL).
 FT DOMAIN 229 248 7 (POTENTIAL).
 FT TRANSMEM 249 270 8 (POTENTIAL).
 FT DOMAIN 271 291 9 (POTENTIAL).
 FT TRANSMEM 292 312 10 (POTENTIAL).
 FT DOMAIN 313 333 11 (POTENTIAL).
 FT TRANSMEM 334 353 12 (POTENTIAL).
 FT DISULFID 116 193 BY SIMILARITY.
 FT CARBOHYD 19 19 POTENTIAL.
 FT NON_TER 353 353
 SO SEQUENCE 353 AA; 39998 MW; F4564B58 CRC32;

Query Match 2.8%; Score 10; DB 1; Length 353;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 Db 140 DRYLAIVHAT 149

RESULT 22
 ID IL8B_RABIT STANDARD; PRT; 358 AA.
 AC P35344;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (GRO/MGSA
 DE RECEPTOR).
 GN IL8RB OR CXCR2.
 OS ORYCTOLAGUS Cuniculus (RABBIT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC LAGOMORPHA; LEPORIDAE; ORYCTOLAGUS.
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-ALBINO; TISSUE-BLOOD;
 RX MEDLINE: 94230294.
 RA PRADO G.N., THOMAS K.M., SUZUKI H., LAROSA G.J., WILKINSON N.C.,
 RA FOLCO E., NAVARRO J.:
 RT "Molecular characterization of a novel rabbit interleukin-8 receptor
 RT isotype".
 RL J. Biol. Chem. 269:12391-12394(1994).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THE AFFINITY OF THIS RECEPTOR IS IL-8 >> NAP-2 >
 CC MGSA (GRO).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREFERENTIALLY IN NEUTROPHILS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L24445; G437662; -.
 DR PIR: A53752; A53752.
 DR GCRDB: GCR_0861; -.
 DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 46 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 47 73 1 (POTENTIAL).
 FT DOMAIN 74 82 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 83 103 2 (POTENTIAL).
 FT DOMAIN 104 118 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 119 140 3 (POTENTIAL).
 FT DOMAIN 141 161 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 162 181 4 (POTENTIAL).
 FT DOMAIN 182 206 5 (POTENTIAL).
 FT TRANSMEM 207 229 6 (POTENTIAL).
 FT DOMAIN 230 249 7 (POTENTIAL).
 FT TRANSMEM 250 271 8 (POTENTIAL).
 FT DOMAIN 272 292 9 (POTENTIAL).
 FT TRANSMEM 293 313 10 (POTENTIAL).
 FT DOMAIN 314 333 11 (POTENTIAL).
 FT CARBOHYD 8 8 POTENTIAL.
 FT DISULFID 117 194 BY SIMILARITY.
 SO SEQUENCE 358 AA; 40632 MW; 90C46E08 CRC32;

Query Match 2.8%; Score 10; DB 1; Length 358;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
 Db 141 DRYLAIVHAT 150

RESULT 23
 ID IL8B_RAT STANDARD; PRT; 359 AA.
 AC P35407;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE HIGH AFFINITY INTERLEUKIN-8 RECEPTOR B (IL-8R B) (CXCR-2) (GRO/MGSA
 DE RECEPTOR).
 GN IL8RB OR CXCR2.
 OS RATTUS NORVEGICUS (RAT).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIOGOGNATHI; MORIDAE; MURINAE; RATTUS.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;
 RA GOBL. A.E., WANG S., ZHOU Y., OEBERG K.;
 RL SUBMITTED (FEB-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=LIVER;
 RA KUNISHI K., SHIBATA F., WATANABE K., TSURUFUJI S., NAKAGAWA H.,
 RA FUJIOKA M.;
 RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=SPLEEN, AND LUNG;
 RA MEDLINE: 97115810.
 RX DUNSTAN C., A.N., SALAFRANCA M.N., ADHIKARI S., XIA Y., FENG L.,
 RA HARRISON J.K.;
 RT Identification of two rat genes orthologous to the human
 RT interleukin-8 receptors.;
 RL J. BIOL. CHEM. 271:32770-32776(1996).
 CC -1- FUNCTION: RECEPTOR TO INTERLEUKIN-8, WHICH IS A POWERFUL
 CC NEUTROPHILS CHEMOTACTIC FACTOR. BINDING OF IL-8 TO THE RECEPTOR
 CC CAUSES ACTIVATION OF NEUTROPHILS. THIS RESPONSE IS MEDIATED VIA A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM. THIS RECEPTOR BINDS TO IL-8 WITH A HIGH AFFINITY
 CC AND TO GRO/MGSA AND NAP-2 ALSO WITH A HIGH AFFINITY.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL: X77797; G498703; -
 DR EMBL: D63584; G944819; -
 DR EMBL: U70988; G1617613; -
 DR PIR: S42096; S42096.
 DR GCRDB: GCR_0913; -
 DR GCRDB: GCR_1403; -
 DR GCRDB: GCR_1524; -
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW CHEMOTAXIS.
 FT DOMAIN 1 47
 FT TRANSMEM 48 74
 FT DOMAIN 49 75
 FT TRANSMEM 75 83
 FT DOMAIN 84 104
 FT TRANSMEM 105 119
 FT DOMAIN 120 141
 FT TRANSMEM 142 162
 FT DOMAIN 163 182
 FT TRANSMEM 183 207
 FT DOMAIN 208 230
 FT TRANSMEM 231 250
 FT DOMAIN 251 272
 FT TRANSMEM 273 296
 FT DOMAIN 297 314
 FT TRANSMEM 315 358
 FT DOMAIN 359 370
 FT TRANSMEM 370 405
 FT CARBOHYD 201 201
 FT CARBOHYD 202 202
 FT DISULFID 118 195
 FT SEQUENCE 359 AA; 40532 MW; 5829D194 CRC32;

Query Match 2.88; Score 10; DB 1; Length 359;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 133 DRYLAIVHAT 142
 |||||||||
 Db 142 DRYLAIVHAT 151
 RESULT 24
 OPRX_CAVPO STANDARD; PRT; 370 AA.
 ID OPRX_CAVPO
 AC P47748;
 DT 01-FEB-1996 (REL. 33, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE NOCICEPTIN RECEPTOR (ORPHANIN FQ RECEPTOR) (KAPPA-TYPE 3 OPIOID
 DE RECEPTOR (KOR-3) (XOR).
 GN OPR1 OR OOR.
 OS CAVIA PORCELLUS (GUINEA PIG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; HYSTRICOGNATHI; CAVIIDAE; CAVIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HARTLEY; TISSUE=BRAIN.
 RA XIE G.;
 RL SUBMITTED (DEC-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: RECEPTOR FOR THE NEUROPEPTIDE NOCICEPTIN/ORPHANIN FQ.
 CC HAS A POTENTIAL ROLE IN MODULATING A NUMBER OF BRAIN FUNCTIONS,
 CC INCLUDING INSTINCTIVE BEHAVIORS AND EMOTIONS. THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLYL
 CC CYCLASE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL: U04369; G606789; -
 DR GCRDB: GCR_1169; -
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 DR PFAM: PF00001; 7tm_1; 1.
 DR HSP: P34996; 1DD.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KW PHOSPHORYLATION; LIPOPROTEIN; PALMITATE.
 FT DOMAIN 1 50
 FT TRANSMEM 51 77
 FT DOMAIN 78 87
 FT TRANSMEM 88 109
 FT DOMAIN 110 124
 FT TRANSMEM 125 146
 FT DOMAIN 147 165
 FT TRANSMEM 166 188
 FT DOMAIN 189 211
 FT TRANSMEM 212 236
 FT DOMAIN 237 264
 FT TRANSMEM 265 288
 FT DOMAIN 289 300
 FT TRANSMEM 301 322
 FT DOMAIN 323 370
 FT TRANSMEM 370 405
 FT CARBOHYD 201 201
 FT CARBOHYD 202 202
 FT DISULFID 118 195
 FT SEQUENCE 370 AA; 40789 MW; 4ADFAEEF CRC32;

RE J. NEUROINMUNOL. 59:91-101(1995).
CC -1- FUNCTION: RECEPTOR FOR THE NEUROPEPTIDE NOCICEPTIN/ORPHANIN FQ.
CC HAS A POTENTIAL ROLE IN MODULATING A NUMBER OF BRAIN FUNCTIONS,
CC INCLUDING INSTINCTIVE BEHAVIORS AND EMOTIONS. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLYL
CC CYCLASE.
CC
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U04952; G440880; -
CC EMBL: D31657; G808874; -
CC EMBL: D31656; G808874; JOINED.
CC EMBL: X91813; G1008982; -
CC EMBL: U32932; G1464791; -
CC EMBL: U32928; G1464791; JOINED.
CC EMBL: U32930; G1464791; JOINED.
CC EMBL: U09421; G551485; -
CC EMBL: U14165; G540093; -
CC GCRDB: GCR_0891; -
CC GCRDB: GCR_1715; -
CC GCRDB: GCR_1715; -
CC GCRDB: GCR_1728; -
CC GCRDB: GCR_1731; -
CC MGD: MGI:97440; OPRL.
CC PROSITE: PS00237; G-PROTEIN-RECEPTOR: 1.
CC PFM: PFM0001; 7em_1; 1.
CC HSSP: P34966; 1DDO.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW PHOSPHORYLATION; LIPOPROTEIN; PALMITATE.
FT DOMAIN 1 47
FT TRANSMEM 48 74
FT DOMAIN 75 84
FT TRANSMEM 85 106
FT DOMAIN 107 121
FT TRANSMEM 122 143
FT DOMAIN 144 162
FT TRANSMEM 163 185
FT DOMAIN 186 208
FT TRANSMEM 209 233
FT DOMAIN 234 261
FT TRANSMEM 262 285
FT DOMAIN 286 297
FT TRANSMEM 298 319
FT DOMAIN 320 365
FT DISULFID 120 197
FT LIPID 331 331
FT CARBOHYD 21 21
FT CARBOHYD 26 26
FT CARBOHYD 36 36
FT CONFLICT 348 349
FT SEQUENCE 367 AA: 40491 MW: 35472156 CRC32;
SQ

Query Match 2.8%; Score 10; DB 1; Length 367;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 CLNPITLYAFL 305
DB 310 CLNPITLYAFL 319

RESULT 27
OPRX_PIG STANDARD: PRT: 370 AA.

AC P79292;
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE NOCICEPTIN RECEPTOR (ORPHANIN FQ RECEPTOR) (KAPPA-TYPE 3 OPIOID
DE RECEPTOR) (KOR-3) (ORGC) (K3 OPIATE RECEPTOR).
GN OPRL.
OS SCS SCROFA (PIG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN CORTEX;
RA OSINSKI M.A., PAMPUSCH M.S., BROWN D.R., MORTANGH M.P.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDJ DATA BANKS.
CC -1- FUNCTION: RECEPTOR FOR THE NEUROPEPTIDE NOCICEPTIN/ORPHANIN FQ.
CC HAS A POTENTIAL ROLE IN MODULATING A NUMBER OF BRAIN FUNCTIONS,
CC INCLUDING INSTINCTIVE BEHAVIORS AND EMOTIONS. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLYL
CC CYCLASE (BY SIMILARITY).
CC
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U72758; G1763257; -
CC GCRDB: GCR_1519; -
CC PROSITE: PS00237; G-PROTEIN-RECEPTOR: 1.
CC PFM: PFM0001; 7em_1; 1.
CC HSSP: P34966; 1DDO.
KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
KW PHOSPHORYLATION; LIPOPROTEIN; PALMITATE.
FT DOMAIN 1 50
FT TRANSMEM 51 77
FT DOMAIN 78 87
FT TRANSMEM 88 109
FT DOMAIN 110 124
FT TRANSMEM 125 146
FT DOMAIN 147 165
FT TRANSMEM 166 188
FT DOMAIN 189 211
FT TRANSMEM 212 236
FT DOMAIN 237 264
FT TRANSMEM 265 288
FT DOMAIN 289 300
FT TRANSMEM 301 322
FT DOMAIN 323 370
FT DISULFID 123 200
FT LIPID 334 334
FT CARBOHYD 21 21
FT CARBOHYD 28 28
FT CARBOHYD 39 39
FT SEQUENCE 370 AA: 40610 MW: E1050DAC CRC32;
SQ

Query Match 2.8%; Score 10; DB 1; Length 370;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 CLNPITLYAFL 305
DB 313 CLNPITLYAFL 322

RESULT 28
OPRX_RAT STANDARD: PRT: 367 AA.

AC P35370:
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE NOCICEPTIN RECEPTOR (ORPHANIN FQ RECEPTOR) (KAPPA-TYPE 3 OPIOID
 RECEPTOR) (KOR-3) (FOR-C) (XOR1).
 GN OPRI1 OR OOR.
 OS RATTUS NORVEGICUS (RAT).
 OC EURASYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIUROGNATHI; MORIDAE; MORINAE; RATTUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MISTAR; TISSUE=BRAIN;
 RX MEDLINE: 94215703.
 RA FUKUDA K., KATO S., MORI K., NISHI M., TAKESHIMA H., INABE N.,
 RA MIYATA T., HOUTANI T., SUGIMOTO T.,
 RT "CDNA cloning and regional distribution of a novel member of the
 RT opioid receptor family.";
 RL FEBS LETT. 343:42-46(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HIPPOCAMPUS;
 RA MENG F., XIE G., ALFRED M., THOMPSON R., HOVERSTEN M., WATSON S.,
 RA AKIL H.;
 RL SUBMITTED (JAN-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
 RX MEDLINE: 94307401.
 RA BUNZOW J.R., SAEZ C., MORTIUD M., BOUVIER C., WILLIAMS J.T., LOW M.,
 RA GRANDY D.K.;
 RT "Molecular cloning and tissue distribution of a putative member of
 RT the rat opioid receptor gene family that is not a mu, delta or kappa
 RT opioid receptor type.";
 RL FEBS LETT. 347:284-288(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC MEDLINE: 94307400.
 RA CHEN Y., FAN Y., LU J., MESTEK A., TIAN M., KOZAK C.A., YU L.;
 RT "Molecular cloning, tissue distribution and chromosomal localization
 RT of a novel member of the opioid receptor gene family.";
 RL FEBS LETT. 347:279-283(1994).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE: 95096849.
 RA LACHOWICZ J.E., SHEN Y., MONSMA F.J. JR., SIBLEY D.R.;
 RT "Molecular cloning of a novel G protein-coupled receptor related to
 RT the opiate receptor family.";
 RL J. NEUROCHEM. 64:34-40(1995).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE: 94298959.
 RA WANG J.B., JOHNSON P.S., IMAI Y., PERSICO A.M., OZENBERGER B.A.,
 RA EPPLER C.M., UHL G.R.;
 RT "CDNA cloning of an orphan opiate receptor gene family member and its
 RT splice variant.";
 RL FEBS LETT. 348:75-79(1994).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
 RX MEDLINE: 95182817.
 RA WICK M.J., MINNERATH S.R., LIN X., ELDE R.P., LAW P.Y., LOH H.H.;
 RT "Isolation of a novel cDNA encoding a putative membrane receptor with
 RT high homology to the cloned mu, delta, and kappa opioid receptors.";
 RL BRAIN RES. MOL. BRAIN RES. 27:37-44(1994).
 CC -1- FUNCTION: RECEPTOR FOR THE NEUROPEPTIDE NOCICEPTIN/ORPHANIN FQ.
 CC HAS A POTENTIAL ROLE IN MODULATING A NUMBER OF BRAIN FUNCTIONS,
 CC INCLUDING INSTINCTIVE BEHAVIORS AND EMOTIONS. THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH INHIBITS ADENYLYL
 CC CYCLASE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SEVERAL BRAIN AREAS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D16438; G533355; -
 DR EMBL: U05239; G451844; -
 DR EMBL: U01913; G487965; -
 DR EMBL: L28144; G496220; -
 DR EMBL: U07871; G606803; -
 DR EMBL: L33916; G557200; -
 DR EMBL: L29419; G510719; -
 DR PIR: S46238; S46238.
 DR PIR: S43655; S43655.
 DR GCRDB: GCR_0834; -
 DR GCRDB: GCR_0898; -
 DR GCRDB: GCR_0912; -
 DR GCRDB: GCR_1030; -
 DR GCRDB: GCR_1455; -
 DR GCRDB: GCR_1487; -
 DR GCRDB: GCR_1517; -
 DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
 DR PRAM: PF00001; 7tm_1; 1.
 DR HSP: P34996; 1DD0.
 KW G-PROTEIN COUPLED RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN;
 KM PHOSPHORYLATION; LIPOPROTEIN; PALMITATE.
 FT DOMAIN 1 47
 FT TRANSMEM 48 74
 FT DOMAIN 75 84
 FT TRANSMEM 85 106
 FT DOMAIN 107 121
 FT TRANSMEM 122 143
 FT DOMAIN 144 162
 FT TRANSMEM 163 185
 FT DOMAIN 186 208
 FT TRANSMEM 209 233
 FT DOMAIN 234 261
 FT TRANSMEM 262 285
 FT DOMAIN 286 297
 FT TRANSMEM 298 319
 FT DOMAIN 320 366
 FT DISULFID 120 197
 FT LIPID 331 331
 FT CARBOHYD 21 21
 FT CARBOHYD 26 26
 FT CARBOHYD 36 36
 FT CONFLICT 105 105
 FT CONFLICT 105 105
 FT CONFLICT 226 226
 FT CONFLICT 246 246
 FT CONFLICT 348 348
 SQ SEQUENCE 367 AA; 40523 MW; 7FD40CCC CRC32;

Query Match 2.88; Score 10; DB 1; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.025;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Gy 296 CLNPITVAF 305
 |||||
 Db 310 CLNPITVAF 319

Search completed: August 8, 1999, 06:07:34
 Job time: 501 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 8, 1999, 02:04:38 ; Search time 17.77 Seconds
(without alignments)
1219.097 Million cell updates/sec

Title: US-09-104-063-4
Perfect score: 352
Sequence: 1 MEGSITSDNYTEMGSGD.....KRGHSSVSTESSSFHSS 352

Scoring table: OLIGO

Searched: 201082 seqs, 61543640 residues

Database :

SPTREMBL_10.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_undefined:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	347	98.6	350	060835	060835 homo sapien
2	176	50.0	352	6	077488
3	60	17.0	332	11	062973
4	39	11.1	358	13	09YGC3
5	19	5.4	353	13	093247
6	16	4.5	357	13	042445
7	10	2.8	356	6	097571

ALIGNMENTS

RESULT 1
060835 PRELIMINARY: PRT: 360 AA.
AC 060835:
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE CXCR4 GENE ENCODING RECEPTOR CXCR4.
CN CXCR4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homidae; Homo.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 98258970.
RA CARUZ A., SANSOM M., ALONSO J.M., ALCAMI J., BALEUX F.,
VIRELITIER J.L., PARMENTIER M., ARENZANA-SEISDEDOS F.;

RT "Genomic organization and promoter characterization of human CXCR4
RT gene".
RL FEBS Lett. 426:271-278(1998).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AJ224869; CAA12166.1; -
DR PFM: PF00001; 7tm.1; 1
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
SQ SEQUENCE 360 AA; 40607 MW; 110EB9E6 CRC32;

Query Match 98.6%; Score 347; DB 4; Length 360;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 IYTSNDYTEMGSGDYSKPECFREENANFNKIFLPTIYSIIFLTGIYNGILVMGY 65
DB 14 IYTSNDYTEMGSGDYSKPECFREENANFNKIFLPTIYSIIFLTGIYNGILVMGY 73
QY 66 QKKLSMTDKYRLHLSVADLLFVITLPEWADAVANWYGNFLCAVHYIVNLYSSVL 125
DB 74 QKKLSMTDKYRLHLSVADLLFVITLPEWADAVANWYGNFLCAVHYIVNLYSSVL 133
QY 126 ILAFISLDRYLAIVHATNSQRPRLKLAEKVYVGVWIPALLITIDPFIANNSEADRYI 185
DB 134 ILAFISLDRYLAIVHATNSQRPRLKLAEKVYVGVWIPALLITIDPFIANNSEADRYI 193
QY 186 CDREFPNDLMVYVFOHIMVGLIPGIVILSCYCIISKLSHSGHOKRKALKTIVILI 245
DB 194 CDREFPNDLMVYVFOHIMVGLIPGIVILSCYCIISKLSHSGHOKRKALKTIVILI 253
QY 246 LAFPCMLPYIIGISIDSTILLEITIKGCEFEFNTYHKMISTEALAFPCCLNPILYATL 305
DB 254 LAFPCMLPYIIGISIDSTILLEITIKGCEFEFNTYHKMISTEALAFPCCLNPILYATL 313
QY 306 GAKFITSAGHALTSVRSGLKILSKRGHSSVSTESSSFHSS 352
DB 314 GAKFITSAGHALTSVRSGLKILSKRGHSSVSTESSSFHSS 360

RESULT 2
077488 PRELIMINARY: PRT: 352 AA.
ID 077488:
AC 077488:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE CXCR4 RECEPTOR.
OS Cercopithecus aethiops (Green monkey) (Givet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae;
OC Chlorocebus.
RN (1)
RP SEQUENCE FROM N.A.
RA MURAYAMA Y., MATSUNAGA S., INOUE-MURAYAMA M.:
RT "cDNA sequence of African green monkey CXCR-4 chemokine receptor
RT gene".
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC EMBL: AB015943; BAA31327.1; -
DR PFM: PF00001; 7tm.1; 1.
DR PROSITE: PS00237; G-PROTEIN_RECEPTOR; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
SQ SEQUENCE 352 AA; 39773 MW; BD6D5150 CRC32;

Query Match 50.0%; Score 176; DB 6; Length 352;
Best Local Similarity 100.0%; Pred. No. 7; 6e-168;
Matches 176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 177 VSEADRYICDRFPNDLMVYVFOHIMVGLIPGIVILSCYCIISKLSHSGHOKRK 236

```

|||||
Db 177 VSEADRYICDFEYNNDLVVVFQFQIHWGLIPGIVILSCYCIILSKLSKSHQKRC 236
OY 237 AKTIVILLAFACWMLPYIGISIDSFILFLIKOGCFEFTVHKWISITLALAFHCC 296
Db 237 AKTIVILLAFACWMLPYIGISIDSFILFLIKOGCFEFTVHKWISITLALAFHCC 296
OY 237 LNPILYAFGAFKTSQAQHALTSVSRGSSLYKLKSKGRGSHSVSTESSESSPHSS 352
Db 237 LNPILYAFGAFKTSQAQHALTSVSRGSSLYKLKSKGRGSHSVSTESSESSPHSS 352

```

```

RESULT 3
ID 062973 PRELIMINARY; PRT: 332 AA.
AC 062973;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE CHEMOKINE RECEPTOR LCRI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA WONG M., XIN W.W., DOMAN R.S.;
RL MOL. PSYCH. 0:0-0(0).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: U54791; AAB01981.1; -.
DR PFAM: PF00001; 7tm_1; 2.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein.
KW SEQUENCE 332 AA; 37442 MW; C0EAB84B CRC32;

```

```

Query Match 17.0%; Score 60; DB 11; Length 332;
Best Local Similarity 100.0%; Pred. No. 7.1e-52;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 197 VWFQFQIHWGLIPGIVILSCYCIILSKLSKSHQKRRALKTTVILLAFACWMLPY 256
Db 177 VWFQFQIHWGLIPGIVILSCYCIILSKLSKSHQKRRALKTTVILLAFACWMLPY 236

```

```

RESULT 4
ID 09YGC3 PRELIMINARY; PRT: 358 AA.
AC 09YGC3;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE CHEMOKINE RECEPTOR 4.
GN CXCR4.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia; Anura;
OC Mesobatrachia; Pipidae; Pipidae; Xenopodidae; Xenopus.
RN
RP SEQUENCE FROM N.A.
RA MOEPPS B., KNOEFFLE K., BROWN M., KNOEHEL W., GIERSCHEK P.;
RT "Expression of the CXCR chemokine receptor 4 during early Xenopus
RT leavis embryogenesis: a possible role of chemokine receptors as
RT regulators of development and differentiation.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: Y17895; CAAT6924.1; -.
DR EMBL: Y17894; CAAT6923.1; -.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein.
KW SEQUENCE 358 AA; 40078 MW; E59JFDID CRC32;

```

```

Query Match 11.1%; Score 39; DB 13; Length 358;
Best Local Similarity 100.0%; Pred. No. 7.6e-31;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 107 FLCKAVHYIYTNLYSVLLAFISLDRLYLAIVHATNSQ 145
Db 111 FLCKAVHYIYTNLYSVLLAFISLDRLYLAIVHATNSQ 149

```

```

RESULT 5
ID 093247 PRELIMINARY; PRT: 353 AA.
AC 093247;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE CXCR4.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinidae; Cyprininae; Cyprinus.
RN
RP SEQUENCE FROM N.A.
RA FUJIKI K., NAKAO M., SHIN D., YANO T.;
RT "CDNA cloning of a carp homologue of mammalian CXCR4.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AB012310; BAA32797.1; -.
DR PFAM: PF00001; 7tm_1; 1.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein.
KW SEQUENCE 353 AA; 39633 MW; 23DD5347 CRC32;

```

```

Query Match 5.4%; Score 19; DB 13; Length 353;
Best Local Similarity 100.0%; Pred. No. 7.5e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 119 NLXSVLLAFISLDRLYLA 137
Db 117 NLXSVLLAFISLDRLYLA 135

```

```

RESULT 6
ID 042445 PRELIMINARY; PRT: 357 AA.
AC 042445;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE CXCR4 CHEMOKINE RECEPTOR.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
OC Neopterygii; Teleostei; Euteleostei; Protacanthopterygii;
OC Salmoniformes; Salmonidae; Oncorhynchus.
RN
RP SEQUENCE FROM N.A.
RA DANIELS G.D., CHARLEMAGNE J., SECOMBES C.J.;
RT "Cloning and sequencing of a rainbow trout, Oncorhynchus mykiss,
RT chemokine receptor homolog.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL: AJ001039; CA04493.1; -.
DR PFAM: PF00001; 7tm_1; 1.
DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein.
KW SEQUENCE 357 AA; 39817 MW; 6D892EFE CRC32;

```

```

Query Match 4.5%; Score 16; DB 13; Length 357;

```

Best Local Similarity 100.0%; Pred. No. 7.5e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 73 TDYRLHLSVADLFFV 88
|||||
Db 75 TDYRLHLSVADLFFV 90

RESULT 7

097571
ID 097571 PRELIMINARY; PRT; 356 AA.
AC 097571;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
DE INTERLEUKIN-8 RECEPTOR.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BEED BEAGLE;
RA CHANG Y.F., NOVOSIL V., CHANG C.F.;
RT "The isolation and sequence of canine interleukin-8 receptor
homolog."
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF047047; AAC98968.1; -
KW Receptor
SQ SEQUENCE 356 AA; 40505 MW; C2B7961F CRC32;

Query Match 2.8%; Score 10; DB 6; Length 356;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 133 DRYLAIVHAT 142
|||||
Db 141 DRYLAIVHAT 150

Search completed: August 8, 1999, 02:45:06
Job time: 2428 sec

